USSN 09/719,277

JOURNAL OF VIROLOGY, Jan. 2003, p. 1578–1583
0022-538X/03/$08.00+0 DOI: 10.1128/JVI.77.2.1578–1583.2003

Vol. 77, No. 2

Hepatitis C Virus F Protein Is a Short-Lived Protein Associated with the Endoplasmic Reticulum

Zhenming Xu, Jinah Choi, Wen Lu, and Jing-hsiung Ou*

Department of Molecular Microbiology and Immunology, University of Southern California, Keck School of Medicine, Los Angeles, California 90033-1054

Received 5 September 2002/Accepted 17 October 2002

Hepatitis C virus (HCV) F protein is a newly discovered HCV gene product that is expressed by translational ribosomal frameshift. Little is known about the biological properties of this protein. By performing pulse-chase labeling experiments, we demonstrate here that the F protein is a labile protein with a half-life of <10 min in Huh7 hepatoma cells and in vitro. The half-life of the F protein could be substantially increased by proteasome inhibitors, suggesting that the rapid degradation of the F protein is mediated by the proteasome pathway. Further immunofluorescence staining and subcellular fractionation experiments indicate that the F protein is primarily associated with the endoplasmic reticulum. This subcellular localization is similar to those of HCV core and NS5A proteins, raising the possibility that the F protein may participate in HCV morphogenesis or replication.

Hepatitis C virus (HCV) can cause severe liver diseases, including hepatitis, liver cirrhosis, and hepatocellular carcinoma. This virus belongs to the flavivirus family and has a positive-stranded RNA genome of ca. 9.6 kb. The genome of this virus contains a long open reading frame, which codes for a polyprotein with a length of slightly more than 3,000 amino acids (2). This polyprotein is cleaved by cellular and viral proteases to generate 10 viral gene products: core, E1, E2, p7, NS2, NS3, NS4A, NS4B, NS5A, and NS5B. The core protein is the viral capsid protein with a length of 191 amino acids (p21c). It can be further cleaved to generate a smaller 179-amino-acid core protein (p19c) (15). E1 and E2 are the two viral envelope proteins. p7 contains two transmembrane domains (4), and NS2 likely contains at least four transmembrane domains (29). The functions of these two proteins are not clear. NS3 is a protease and a helicase. NS4A is a cofactor for the NS3 protease, and NS5B is the viral RNA polymerase. The functions of NS4B and NS5A are not totally clear, but they are likely involved in viral RNA replication and pathogenesis. All of these HCV proteins are believed to form replication complexes on intracellular membranes for either viral morphogenesis or RNA replication (7, 16, 23). The translation of the HCV polyprotein sequence is mediated by an internal ribosomal entry site, which comprises most of the 5′-noncoding region and the first few codons of the polyprotein coding sequence (17).

In addition to the above 10 gene products, a new HCV protein named F protein was also recently reported (24, 25, 28). This protein is apparently expressed during natural HCV infection, since its reactive antibodies were detected in HCV patients. The F protein is encoded by a reading frame that overlaps the core protein coding sequence (Fig. 1A). This F protein is expressed by a −2/+1 ribosomal frameshift during translation (24, 28). Based on the radiosequencing of the F protein synthesized in vitro, the ribosomal frameshift site for the synthesis of the F protein is located in the A-rich sequence at codons 10 to 12 of the core protein sequence (28). Thus, the F protein and the core protein have the same amino-terminal sequence. Their sequences diverge after 10 amino acids. The length of the F protein varies depending on the genotypes. For genotype 1a such as the HCV-1 isolate, the F protein is 161 amino acids long.

In an attempt to understand the biological functions of the F protein, we have expressed the F protein of the HCV-1 isolate in Huh7 cells, a well-differentiated human hepatoma cell line. To facilitate the analysis, a single nucleotide was deleted from the stretch of 10 adenosines located at codons 8 to 11 (Fig. 1A). This nucleotide deletion fused the first 10 codons of the core protein to the −2/+1 reading frame. Hence, the predominant protein product produced from this sequence would be the F protein (Fig. 1A) (28). Since the HCV ribosomal frameshift signal can also mediate −1 translational frameshift (24; J. Choi, Z. Xu, and J.-H. Ou, unpublished data), which will restore the core protein translation (see below), a C-to-T mutation was also created at nucleotide (nt) 57 to generate a premature termination codon in the core protein coding sequence (Fig. 1A). This additional mutation prevented the expression of the core protein without affecting the F protein coding sequence. The antigenic epitope of hemagglutinin (HA) was fused to the amino terminus of this mutated sequence, which was then inserted into the pCDEF vector under the expression control of the promoter of the elongation factor 1α gene (28). The resulting DNA plasmid, pCDEF-HAF, was then transfected into Huh7 cells for the expression studies. The HA-tagged F protein was metabolically labeled with [^{35}S]methionine, followed by immunoprecipitation with the anti-HA antibody. As shown in Fig. 1B, the HA-tagged F protein could be detected in cells transfected with pCDEF-HAF but not in cells transfected with the control vector pCDEF. Similar re-

* Corresponding author. Mailing address: Department of Molecular Microbiology and Immunology, University of Southern California, Keck School of Medicine, 2011 Zonal Ave., HMR-401, Los Angeles, CA 90033. Phone: (323) 442-1720 Fax: (323) 442-1721. E-mail: jamesou@hsc.usc.edu.

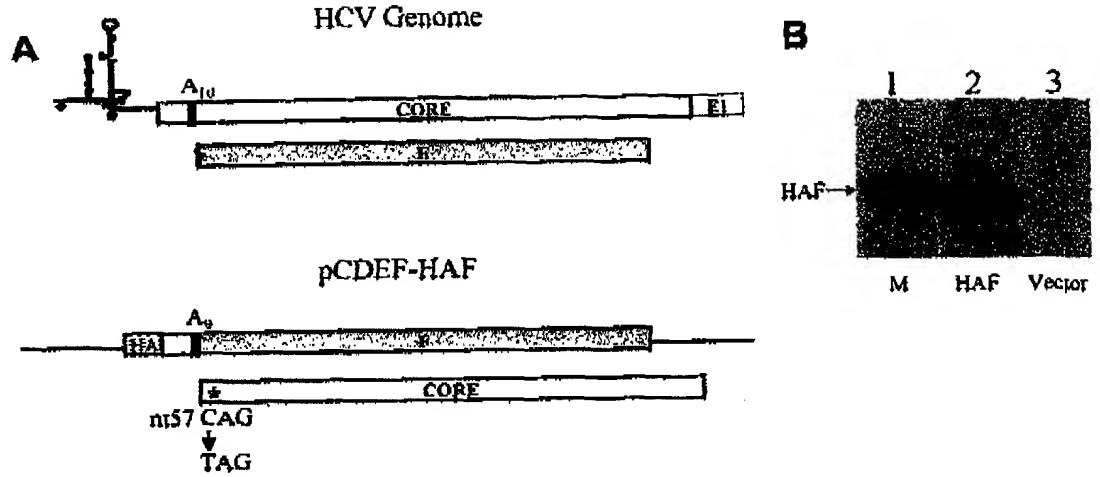

FIG. 1. Expression of the HA-F protein in Huh7 hepatoma cells. (A) Illustrations of the 5′ end of the HCV-1 genome (top) and the HA-F cDNA construct (bottom). The HCV genome shown contains the 5′ noncoding region and the coding regions of the core protein and the F protein. The F protein coding sequence is shaded. A_{10}, the stretch of 10 adenosines at codons 8 to 11 of the core protein coding sequence. This sequence contains the ribosomal frameshift signal for the synthesis of the F protein (28). The HCV-1 genome was used for the construction of the plasmid pCDEF-HAF. In this construct, one adenosine was deleted from A_{10} to generate A_9. This deletion fused the first 10 codons of the core protein to the F protein coding sequence. The HA tag, indicated by a stippled box, was fused to the 5′ end of the coding sequence. The location of the nt 57 C-to-T mutation is indicated by an asterisk. This mutation created a TAG termination codon in the core protein sequence. (B) Immunoprecipitation of the HA-F protein. pCDEF-HAF (lane 2) or the control vector pCDEF (lane 3) was transfected into Huh7 cells by using $CaPO_4$ precipitation procedures (13). Cells were starved for methionine for 3 h at 2 days after transfection and then radiolabeled with [^{35}S]methionine for 1 h, followed by immunoprecipitation with anti-HA antibody (13). The [^{35}S]methionine-labeled HA-F protein synthesized in vitro with rabbit reticulocyte lysates (28) was shown in lane 1 to serve as a marker (M).

sults were obtained when HepG2 cells, a human hepatoblastoma cell line, were used for the expression studies (data not shown).

In our expression studies, we noticed that the F protein was unstable and degraded rapidly after its synthesis. For that reason, we have performed the pulse-chase labeling experiment to determine the half-life of the F protein. In this experiment, the HA-tagged F (HA-F) protein was pulse-labeled with [^{35}S]methionine for 10 min and chased with unlabeled methionine for various lengths of time. As shown in Fig. 2A, the amount of the HA-F protein decreased significantly during the chase and became undetectable after 60 min of chase. Its half-life was determined to be ca. 8 to 10 min based on densitometry (Fig. 2B). In contrast, the amount of the HA-tagged core (HA-core) protein was not apparently reduced by the chase in a similar pulse-chase labeling experiment (Fig. 2C), a finding in agreement with a previous report that the core protein is a stable protein unless it is truncated (22).

The results shown in Fig. 2 demonstrate that the HA-F protein is a labile protein in Huh7 cells. To determine whether the HA-F protein synthesized in vitro is similarly unstable, HA-F protein was synthesized in vitro by using rabbit reticulocyte lysates and then pulse-labeled with [^{35}S]methionine for 10 min. After the translation reaction was stopped with cycloheximide, the F protein was chased for various lengths of time. As shown in Fig. 3A, the amount of HA-F protein was reduced significantly during the chase; these findings are similar to the results observed in Huh7 cells. The half-life was determined to be slightly less than 10 min based on densitometry readings (Fig. 3B). Similar results were obtained when the F protein was expressed in vitro without the HA tag (data not shown).

To understand the molecular mechanism that is responsible for the instability of the F protein, serine protease inhibitors, including leupeptin, aprotinin, and Pefabloc, were added into the translation mixture at the beginning of the chase in separate pulse-chase labeling experiments. None of these protease inhibitors could increase the half-life of the F protein (data not shown). In contrast, if the proteasome inhibitor MG132 was added into the translation mixture at the beginning of the chase, the decrease of the HA-F protein signal was unapparent during the 2-h chase period (Fig. 4A). These results indicated that the degradation of the F protein was most likely mediated by the proteasome pathway in vitro. To determine whether the HA-F protein was also degraded by the same pathway in Huh7 cells, cells transfected with pCDEF-HAF were treated with MG132 or the control solvent dimethyl sulfoxide (DMSO) for 6 h, lysed, and then analyzed by Western blotting with anti-HA antibody. As shown in Fig. 4B, the expression level of the HA-F protein was low in DMSO-treated cells (lanes 1 and 6). However, this expression level was substantially increased if cells were treated with MG132 (lanes 2 and 7). Similar results were obtained when Huh7 cells were treated with lactacystin, another proteasome inhibitor (lane 3). These results indicated that the F protein was likely also degraded by the proteosome pathway in Huh7 cells. Note that in lanes 6 and 7 the cells were transfected with an expression plasmid identical to pCDEF-HAF except that nt 57 was not mutated from C to T, which would create a premature termination codon in the core pro-

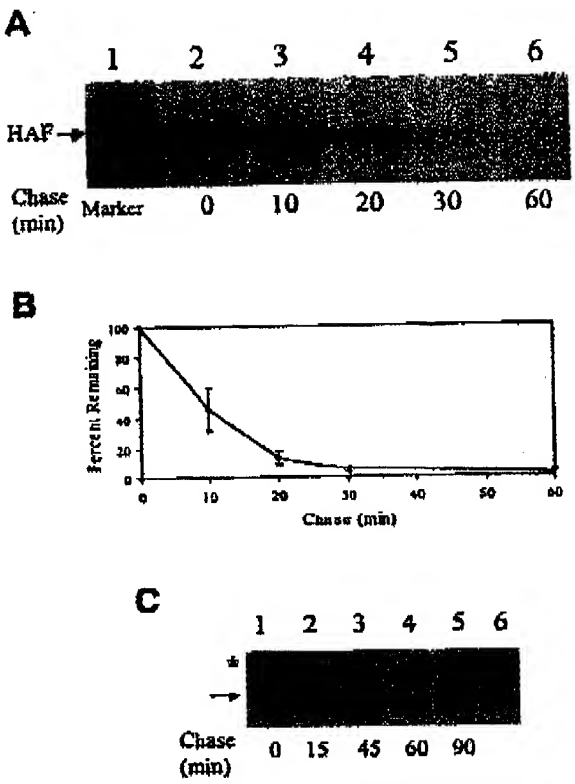

FIG. 2. Pulse-chase labeling experiment of the HCV proteins expressed in Huh7 cells. (A) Pulse-chase labeling experiment of the HA-F protein. Huh7 cells transfected with pCDEF-HAF by $CaPO_4$ precipitation were pulse-labeled with [^{35}S]methionine for 10 min and chased with unlabeled methionine for 0, 10, 20, 30, and 60 min (lanes 2 to 6). Cells were then lysed for immunoprecipitation with the anti-HA antibody by our previous procedures (14). Lane 1 is the [^{35}S]methionine-labeled HA-F protein marker, which was synthesized in vitro by using the rabbit reticulocyte lysates. (B) Determination of the half-life of the HA-F protein. The autoradiogram shown in panel A was analyzed with SigmaScan. The results represented the average of three independent experiments. The HA-F protein level at the zero time point of chase was defined as 100%. (C) Pulse-chase labeling experiment of the HA-core protein. Huh7 cells transfected with pCDEF-HA-core (13) were pulse-labeled with [^{35}S]methionine for 10 min and chased with unlabeled methionine for 0, 15, 45, 60, and 90 min (lanes 1 to 5). Lane 6 was Huh7 cells transfected with the control pCDEF vector. The asterisk marks the location of a nonspecific protein band. The arrow denotes the core protein band.

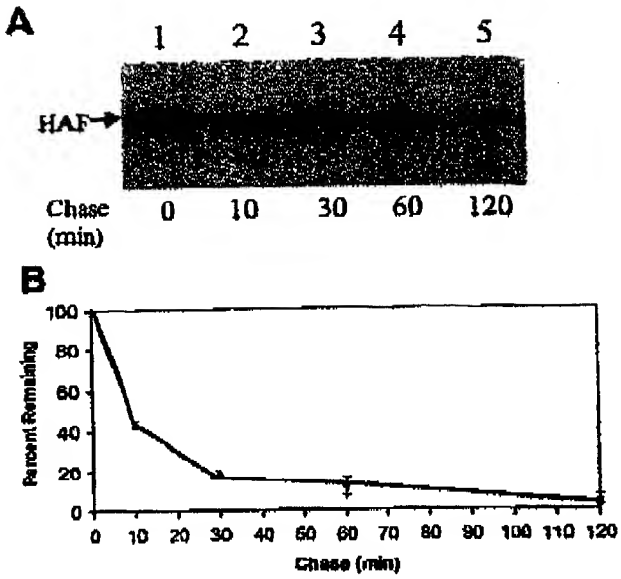

FIG. 3. Pulse-chase labeling experiments of the HA-F protein synthesized in vitro. (A) Autoradiograms of the pulse-chase experiments. The HA-F coding sequence was inserted into pRc/CMV (Invitrogen). The HA-F RNA was then synthesized by using the T7 RNA polymerase and translated with the rabbit reticulocyte lysates. Details of these experimental procedures had been described (28). The HA-F protein was pulse-labeled with [^{35}S]methionine for 10 min. The translation reaction was then stopped by the addition of cycloheximide to a final concentration of 400 μM. The HA-F protein was then chased for 0, 10, 30, 60, or 120 min (lanes 1 to 5). (B) Half-life of the HA-F protein in vitro. The results shown in panel A were quantified with SigmaScan. The results represented the average of three independent experiments. The HA-F protein level at the zero time point was defined as 100%.

tein coding sequence (Fig. 1A). In this case, the HA-tagged core protein was also detected. This result is consistent with the recent findings that the HCV ribosomal frameshift signal could also mediate the −1 ribosomal frameshift to express the core protein from the HAF sequence (24; Choi et al., unpublished).

To investigate the possible biological functions of the F protein, we have also analyzed its subcellular localization. Huh7 cells transfected with pCDEF-HAF were fixed with formaldehyde and stained with the anti-HA antibody. As shown in Fig. 5, the HA-F protein displayed a punctate and perinuclear staining pattern similar to that of calreticulin, a protein associated with the endoplasmic reticulum (ER). This subcellular localization was also similar to that of the core protein in a cotransfection experiment. Stable Huh7 cells that contain bicistronic HCV RNA replicons that contained the neomycin gene and the HCV NS3-NS5 sequence had been produced by several laboratories (1, 8, 9, 11). We had also established stable Huh7 cells containing the HCV replicon (unpublished data). The HA-F protein expressed in these Huh7 cells by transient transfection also displayed a subcellular localization similar to that of NS5A (Fig. 5). Both core and NS5A proteins had previously been shown to localize predominantly to the ER membranes (3, 10, 15, 19, 20). Although the subcellular localization of the HA-F protein was similar to those of core and NS5A, coimmunoprecipitation experiments failed to demonstrate a direct physical interaction between HA-F and the core protein or NS5A (data not shown).

The ER association of the HA-F protein was further supported by the membrane fractionation experiments by using a discontinuous gradient containing 2.0, 1.3, 1.0, and 0.6 M sucrose solutions. The rough ER, the smooth ER, and the Golgi membranes were enriched in the 2.0-1.3, 1.3-1.0, and 1.0-0.6 M interfaces, respectively, in this sucrose gradient (26). To increase the sensitivity of the Western blot analysis, Huh7 cells

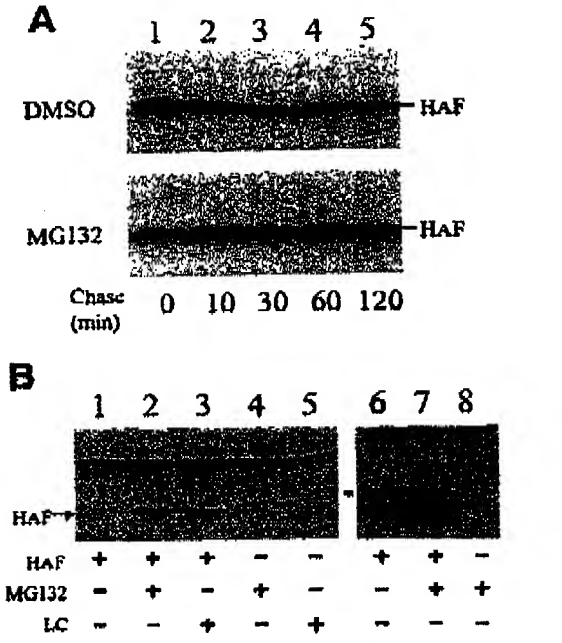

FIG. 4. Stabilization of the HA-F protein by proteasome inhibitors. (A) Pulse-chase labeling experiments of HA-F synthesized in vitro. The HA-F protein was pulse-labeled with [^{35}S]methionine for 10 min as described in the legend to Fig. 3. The translation reactions were then stopped with cycloheximide. MG132 (lower panel) or its control solvent DMSO (upper panel) was then added to a final concentration of 100 μg/ml. The reaction mixtures were chased for 0, 10, 30, 60, or 120 min (lanes 1 to 5). (B) Stabilization of HA-F by proteasome inhibitors in Huh7 cells. pCDEF-HAF (lanes 1 to 3), its control vector pCDEF (lanes 4, 5, and 8), or pCDEF-9aCore (lanes 6 and 7) was transfected into Huh7 cells. pCDEF-9aCore is identical to pCDEF-HAF with the exception that it does not have the C-to-T mutation at nt 57 (28). At 48 h after transfection, cells were treated with DMSO (lane 1), 1 μg of MG132/ml (lanes 2 and 4), or 20 μM of lactacystin (LC) (lanes 3 and 5) for 6 h. Cells were then lysed for Western blot analysis with the anti-HA antibody by using our previous procedures (13). The protein signals were analyzed by the enhanced chemiluminescence kit (Pierce). The expression level of HAF was low in DMSO-treated cells. Hence, lanes 6 to 8 were purposely overexposed on the film to reveal the HAF protein in DMSO-treated cells (lane 6). The asterisk marks the location of the HA-core protein. The identity of this protein was confirmed by Western blotting with the anti-core antibody (data not shown).

were treated with MG132 for 6 h to stabilize the HA-F protein. As shown in Fig. 6, the HA-F protein was detected in both the rough ER and the smooth ER fractions but not the Golgi fraction. To ensure that this gradient indeed faithfully separated ER and Golgi membranes, individual fractions were also analyzed by Western blotting with the anti-GRP78 antibody and wheat germ agglutinin. GRP78 is an ER-associated protein, and wheat germ agglutinin binds specifically to *trans*-Golgi proteins containing clustered terminal *N*-acetylneuraminic acid residues, as well as *N*-acetylglucosamine-containing oligosaccharide chains (26). Calreticulin was not analyzed in these studies due to the lack of a good reactive antibody for Western blot analysis. As shown in Fig. 6, GRP78 was detected mostly in the ER fractions, and proteins with complex-type glycans were detected primarily in the Golgi fraction, indicating that ER and Golgi membranes were faithfully separated in our gradient fractionation procedures. Thus, the results shown in Fig. 6 were in support of the immunofluorescence staining results and indicated that the HA-F protein was primarily an ER-associated protein. The hydrophobicity analysis of the F protein sequence revealed two major hydrophobic domains located at amino acids 28 to 45 and amino acids 95 to 110 (Fig. 7). Either one or both of these two domains may serve as the domains for the F protein to become associated with the ER membranes. Although the length of the HCV F protein varies depending on the genotypes, most of them are longer than 126 amino acids and thus contain these two hydrophobic domains.

In conclusion, we have demonstrated here that the HCV F protein is a short-lived, ER-associated protein. This protein is synthesized by ribosomal frameshift, which occurs at a frequency of ca. 2% in Huh7 cells (28; Choi et al., unpublished). Hence, the finding that this protein is labile with a half-life of <10 min is interesting. It is conceivable that the F protein is needed in only a very small amount in the HCV life cycle. Suzuki et al. recently reported a p17 protein that could be expressed from a genotype 1b HCV core protein coding sequence (22). This protein, which could be detected in their Western blot analysis only when cells were treated with MG132, was thought to be a truncated core protein. Based on our studies, one must carefully reexamine whether the p17 protein detected by these authors was actually the F protein.

It is not unprecedented for an underproduced viral protein to be also a labile protein. The Sindbis virus nsP4 polymerase is expressed by readthrough of an opal codon, followed by proteolytic cleavage, and thus is underproduced compared to other nonstructural proteins (12). This protein is also short-lived and degraded rapidly by the ubiquitin-proteasome N-end rule pathway (6). It has been suggested that the instability of nsP4 may be important for the cessation of minus-strand RNA synthesis during viral replication (6, 21).

The replication of HCV occurs on the membrane structures in the cells, and all of the HCV proteins derived from the polyprotein have been found to associate with the ER membranes either directly or indirectly (4, 5, 7, 15, 18, 27, 29). The finding that the HCV F protein is associated with ER membranes with a subcellular localization similar to those of the HCV core protein and NS5A raises the possibility that the F protein may also be a component of the HCV replication complex. The F protein does not appear to bind directly to the core protein and NS5A. Thus, if it is indeed a component of the HCV replication complex, it will likely interact with this complex through other HCV proteins or indirectly through cellular proteins. The F protein is not needed for HCV RNA replication, because its absence did not impede the replication of the HCV subgenomic RNA replicons (1, 8, 9, 11). However, it remains to be determined whether the F protein may regulate HCV RNA replication or participate in viral morphogenesis. Our finding that the F protein is a short-lived, ER-associated protein will now allow us to further explore its biological functions.

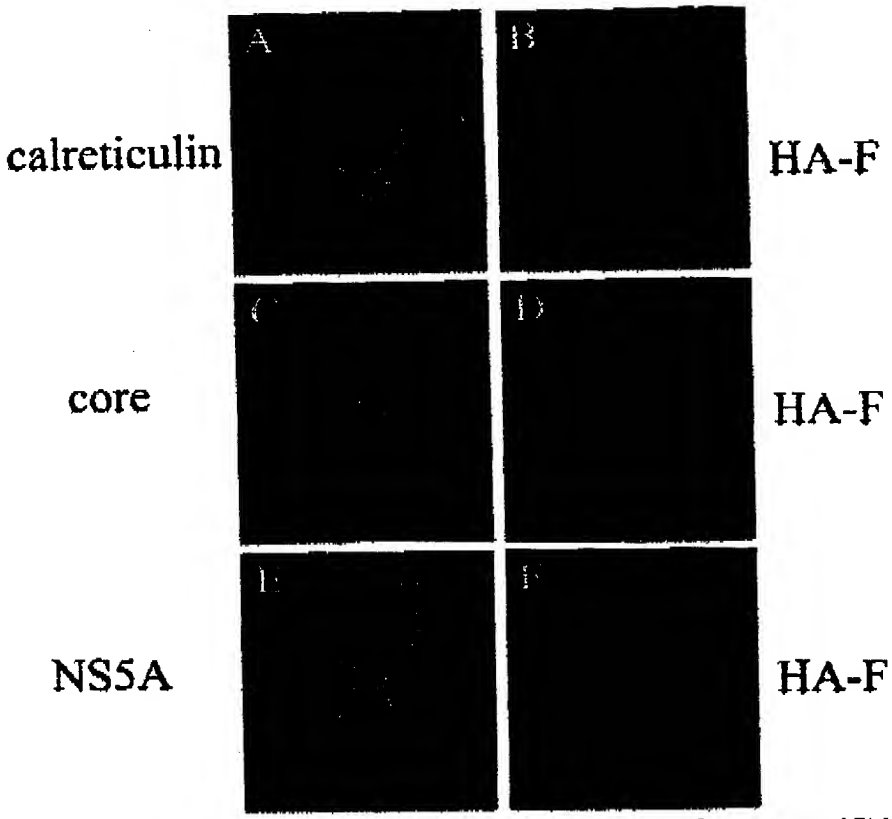

FIG. 5. Immunofluorescence double-staining analysis for the subcellular localization of the HA-F protein. (A and B) Huh7 cells transfected with pCDEF-HAF were double stained with rabbit anti-calreticulin and mouse anti-HA primary antibodies and fluorescein isothiocyanate (FITC)-conjugated goat anti-rabbit and rhodamine (RITC)-conjugated goat anti-mouse secondary antibodies. (C and D) Huh7 cells cotransfected with pCDEF-HAF and pCDEF-core (13) were double stained with rabbit anti-HCV core and mouse anti-HA primary antibodies and FITC-conjugated goat anti-rabbit and RITC-conjugated goat anti-mouse secondary antibodies. (E and F) Huh7 cells containing the HCV subgenomic RNA replicon (unpublished data) were transfected with pCDEF-HAF and double stained with mouse anti-NS5A and rat anti-HA primary antibodies and FITC-conjugated goat anti-mouse and RITC-conjugated goat anti-rat secondary antibodies. In all cases, cells were fixed in 3.7% formaldehyde in phosphate-buffered saline (PBS) at 48 h after transfection for staining by using our previous procedures (14). The HA-F protein was stained in red, whereas all of the other proteins were stained in green. The images were captured with a Nikon confocal microscope at the USC Liver Center.

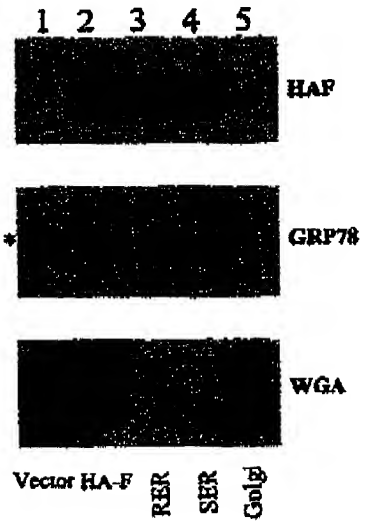

FIG. 6. Membrane fractionation experiments for the analysis of the subcellular localization of HA-F. Huh7 cells transfected with pCDEF-HAF were rinsed with PBS and scraped off the plates into PBS. After a brief centrifugation at 1,500 × *g*, cells were homogenized with a Dounce homogenizer in a 0.25 M sucrose solution containing 10 mM HEPES (pH 7.4), 1 mM EDTA, and 1 mM phenylmethylsulfonyl fluoride. After a brief centrifugation at 15,000 × *g*, the postnuclear supernatant was loaded on a discontinuous sucrose gradient containing 0.6, 1.0, 1.3, and 2.0 M sucrose in 10 mM HEPES (pH 7.4). The gradient was centrifuged at 40,000 rpm by using a Beckman SW40 Ti rotor for 2 h at 4°C as previously described (26). Lane 1, cells transfected with the control vector pCDEF; lanes 2 to 5, cells transfected with pCDEF-HAF; lanes 1 and 2, the postnuclear supernatant prior to fractionation; lane 3, the rough ER (RER) fraction isolated from the 1.3-2.0 M sucrose interface; lane 4, the smooth ER (SER) fraction isolated from the 1.0-1.3 M sucrose interface; lane 5, the Golgi fraction isolated from the 0.6-1.0 M sucrose interface. (Top panel) Western blot analysis with anti-HA antibody; (middle panel) Western blot analysis with anti-GRP78 antibody; (bottom panel) Western blot analysis with horseradish peroxidase-conjugated wheat germ agglutinin (WGA). The asterisk denotes a nonspecific protein band. This protein band did not cofractionate with GRP78 into the ER membrane fractions. Wheat germ agglutinin reacted with multiple protein bands.

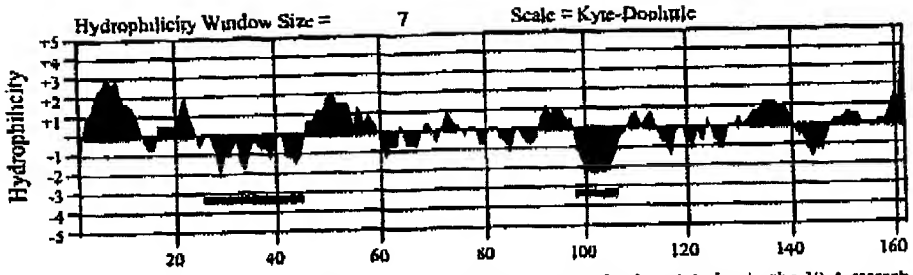

FIG. 7. Hydrophilicity plot of the F protein. The HCV-1 F protein coding sequence with 1 nt deletion in the 10-A stretch was analyzed by the MacVector program. The two thick lines highlight the two major hydrophobic domains.

We thank Michelle Mac Veigh at the USC Liver Center for help with the confocal microscopy.

This work was supported by an American Cancer Society postdoctoral fellowship (PF-01-037-01-MBC) to J.C. and by a research grant from the National Institutes of Health (to J.O.)

REFERENCES

1. **Blight, K. J., A. A. Kolykhalov, and C. M. Rice.** 2000. Efficient initiation of HCV RNA replication in cell culture. Science 290:1972–1974.
2. **Blight, K. J., A. Grahook, H. L. Hanson, and C. M. Rice.** 2002. The molecular biology of hepatitis C virus, p. 81–108. *In* J.-H. J. Ou (ed.), Hepatitis viruses. Kluwer Academic Publishers, Norwell, Mass.
3. **Brass, V., E. Bieck, R. Montserret, B. Wolk, J. A. Hellings, H. E. Blum, F. Penin, and D. Moradpour.** 2002. An amino-terminal amphipathic alpha-helix mediates membrane association of the hepatitis C virus nonstructural protein 5A. J. Biol. Chem. 277:8130–8139.
4. **Carrère-Kremer, S., C. Montpellier-Pala, L. Cocquerel, C. Wychowski, F. Penin, and J. Dubuisson.** 2002. Subcellular localization and topology of the p7 polypeptide of hepatitis C virus. J. Virol. 76:3720–3730.
5. **Cocquerel, L., A. Op de Beeck, M. Lambot, J. Roussel, D. Delgrange, A. Pillez, C. Wychowski, F. Penin, and J. Dubuisson.** 2002. Topological changes in the transmembrane domains of hepatitis C virus envelope glycoproteins. EMBO J. 21:2893–2902.
6. **de Groot, R. J., T. Rumenapf, R. J. Kuhn, E. G. Strauss, and J. H. Strauss.** 1991. Sindbis virus RNA polymerase is degraded by the N-end rule pathway. Proc. Natl. Acad. Sci. USA 88:8967–8971.
7. **Egger, D., B. Wolk, R. Gosert, L. Bianchi, H. E. Blum, D. Moradpour, and K. Bienz.** 2002. Expression of hepatitis C virus proteins induces distinct membrane alterations including a candidate viral replication complex. J. Virol. 76:5974–5984.
8. **Guo, J. T., V. V. Bichko, and C. Seeger.** 2001. Effect of alpha interferon on the hepatitis C virus replicon. J. Virol. 75:8516–8523.
9. **Ikeda, M., M. Yi, K. Li, and S. M. Lemon.** 2002. Selectable subgenomic and genome-length dicistronic RNAs derived from an infectious molecular clone of the HCV-N strain of hepatitis C virus replicate efficiently in cultured Huh7 cells. J. Virol. 76:2997–3006.
10. **Lanford, R. E., L. Notvall, D. Chavez, R. White, G. Frenzel, C. Simonsen, and J. Kim.** 1993. Analysis of hepatitis C virus capsid, E1, and E2/NS1 proteins expressed in insect cells. Virology 197:225–235.
11. **Lohmann, V., F. Korner, J. Koch, U. Herian, L. Theilmann, and R. Bartenschlager.** 1999. Replication of subgenomic hepatitis C virus RNAs in a hepatoma cell line. Science 285:110–113.
12. **Lopez, S., J. R. Bell, E. G. Strauss, and J. H. Strauss.** 1985. The nonstructural proteins of Sindbis virus as studied with an antibody specific for the C terminus of the nonstructural readthrough polyprotein. Virology 141:235–247.
13. **Lu, W., A. Strobecker, and J. H. Ou.** 2001. Post-translational modification of the hepatitis C virus core protein by tissue transglutaminase. J. Biol. Chem. 276:47993–47999.
14. **Lu, W., and J.-H. Ou.** 2003. Phosphorylation of hepatitis C virus core protein by protein kinase A and protein kinase C. Virology 360:20–30.
15. **McLauchlan, J., M. K. Lemberg, G. Hope, and B. Martoglio.** 2002. Intramembrane proteolysis promotes trafficking of hepatitis C virus core protein to lipid droplets. EMBO J. 21:3980–3988.
16. **Mottola, G., G. Cardinali, A. Ceccarelli, C. Trusi, L. Bartholomew, M. R. Torrisi, E. Pedrazzini, S. Bonatti, and G. Migliaccio.** 2002. Hepatitis C virus nonstructural proteins are localized in a modified endoplasmic reticulum of cells expressing viral subgenomic replicons. Virology 293:31–43.
17. **Rijnbrand, R. C., and S. M. Lemon.** 2000. Internal ribosome entry site-mediated translation in hepatitis C virus replication. Curr. Top. Microbiol. Immunol. 242:85–116.
18. **Schmidt-Mende, J., E. Bieck, T. Hugle, F. Penin, C. M. Rice, H. E. Blum, and D. Moradpour.** 2001. Determinants for membrane association of the hepatitis C virus RNA-dependent RNA polymerase. J. Biol. Chem. 276:44052–44063.
19. **Selby, M. J., Q. L. Choo, K. Berger, G. Kuo, E. Glazer, M. Eckart, C. Lee, D. Chien, C. Kuo, and M. Houghton.** 1993. Expression, identification and subcellular localization of the proteins encoded by the hepatitis C viral genome. J. Gen. Virol. 74:1103–1113.
20. **Shi, S. T., S. J. Polyak, H. Tu, D. R. Taylor, D. R. Gretch, and M. M. Lai.** 2002. Hepatitis C virus NS5A colocalizes with the core protein on lipid droplets and interacts with apolipoproteins. Virology 292:198–210.
21. **Sherrin, Y., and J. H. Strauss.** 1994. Regulation of Sindbis virus RNA replication: uncleaved P123 and nsP4 function in minus-strand RNA synthesis, whereas cleaved products from P123 are required for efficient plus-strand RNA synthesis. J. Virol. 68:1874–1885.
22. **Suzuki, R., K. Tamura, J. Li, K. Ishii, Y. Matsuura, T. Miyamura, and T. Suzuki.** 2001. Ubiquitin-mediated degradation of hepatitis C virus core protein is regulated by processing at its carboxyl terminus. Virology 280:301–309.
23. **Tu, H., L. Gao, S. T. Shi, D. R. Taylor, T. Yang, A. K. Mircheff, Y. Wen, A. E. Gorbalenya, S. B. Hwang, and M. M. Lai.** 1999. Hepatitis C virus RNA polymerase and NS5A complex with a SNARE-like protein. Virology 263:30–41.
24. **Varaklioti, A., N. Vassilaki, U. Georgopoulou, and P. Mavromara.** 2002. Alternate translation occurs within the core coding region of the hepatitis C viral genome. J. Biol. Chem. 277:17713–17721.
25. **Walewski, J. L., T. R. Keller, D. D. Stump, and A. D. Branch.** 2001. Evidence for a new hepatitis C virus antigen encoded in an overlapping reading frame. RNA 7:710–721.
26. **Wang, J., A. S. Lee, and J. H. Ou.** 1991. Proteolytic conversion of hepatitis B virus e antigen precursor to end product occurs in a postendoplasmic reticulum compartment. J. Virol. 65:5080–5083.
27. **Wölk, B., D. Sansonno, H. G. Krausslich, F. Dammacco, C. M. Rice, H. E. Blum, and D. Moradpour.** 2000. Subcellular localization, stability, and trans-cleavage competence of the hepatitis C virus NS3-NS4A complex expressed in tetracycline-regulated cell lines. J. Virol. 74:2293–2304.
28. **Xu, Z., J. Choi, T. S. Yen, W. Lu, A. Strobecher, S. Govindarajan, D. Chien, M. J. Selby, and J. Ou.** 2001. Synthesis of a novel hepatitis C virus protein by ribosomal frameshift. EMBO J. 20:3840–3848.
29. **Yamaga, A. K., and J. H. Ou.** 2002. Membrane topology of the hepatitis C virus NS2 protein. J. Biol. Chem. 277:33228–33234.

USSN 09/719,277

Journal of General Virology (2003), 84, 1751–1759 DOI 10.1099/vir.0.19065-0

Characterization of the expression of the hepatitis C virus F protein

Juliette Roussel,[1,2] André Pillez,[1] Claire Montpellier,[1] Gilles Duverlie,[2] Annie Cahour,[3] Jean Dubuisson[1] and Czeslaw Wychowski[1]

[1]CNRS-UPR 2511, IBL/Institut Pasteur de Lille, 59021 Lille Cedex, France

[2]Laboratoire de Virologie, Centre Hospitalier Universitaire-Hôpital Sud, 80054 Amiens Cedex, France

[3]CERVI (Virologie), UPRES EA 2387, Hôpital Pitié-Salpêtrière, 75651 Paris Cedex 13, France

Correspondence
Czeslaw Wychowski
czeslaw.wychowski@ibl.fr

Hepatitis C virus (HCV) is an important human pathogen that affects 170 million people worldwide. The HCV genome is approximately 9·6 kb in length and encodes a polyprotein that is proteolytically cleaved to generate at least 10 mature viral protein products. Recently, a new protein, named F, has been described to be expressed through a ribosomal frameshift within the capsid-encoding sequence, a mechanism unique among members of the family *Flavivirdae*. Here, expression of the F protein was investigated in an *in vitro* transcription/translation assay. Its expression in mammalian cells was confirmed using specific recombinant vaccinia viruses; under these conditions, protein expression is dependent on the HCV IRES. The F protein was tagged with firefly luciferase or the Myc epitope to facilitate its identification. Ribosomal frameshifting was dependent on the presence of mutations in the capsid-encoding sequence. No frameshifting was detected in the absence of any mutation. Furthermore, analysis of the F protein in time-course experiments revealed that the protein is very unstable and that its production can be stabilized by the proteasome inhibitor MG132. Finally, indirect immunofluorescence studies have localized the F protein in the cytoplasm, with notable perinuclear detection.

Received 24 December 2002
Accepted 27 March 2003

INTRODUCTION

Hepatitis C virus (HCV) infection is highly prevalent worldwide and the disease often progresses to chronic hepatitis, leading to cirrhosis and, possibly, hepatocellular carcinoma. Interferon-α (IFN-α) administration in combination with ribavirin remains the most successful treatment for HCV infection but is only effective in about 50% of patients (McHutchison *et al.*, 1998; McHutchison & Poynard, 1999; Schalm *et al.*, 1999). The efficacy of IFN treatment can be limited in duration, leading to the selection of virus variants during the onset of infection that are resistant to IFN.

HCV is related to the flaviviruses and the pestiviruses (Lindenbach & Rice, 2001; Miller & Purcell, 1990; Takeuchi *et al.*, 1990). HCV, a positive-stranded RNA virus with a genomic size of approximately 9·6 kb (Choo *et al.*, 1989, 1991), is now classified within the genus *Hepacivirus*, family *Flaviviridae*. The viral genome contains a large open reading frame encoding a polyprotein of approximately 3000 aa that is cleaved by a combination of host signal peptidases and two virus-encoded proteases to produce the mature structural and non-structural proteins: C, E1, E2, p7, NS2, NS3, NS4A, NS4B, NS5A and NS5B (Lindenbach & Rice, 2001). The C protein, also called the core protein, is the putative capsid protein; E1 and E2 are thought to be the membrane-associated envelope glycoproteins; p7, a polypeptide of unknown function, is cleaved inefficiently from E2; NS2 through NS5B are the non-structural proteins, which, except for NS2, are involved in the replication of the viral genome (Lohmann *et al.*, 1999).

The amino acid sequence of the core protein is well conserved among different HCV isolates. Its N-terminal region is highly basic, while its C-terminal region is hydrophobic (reviewed by McLauchlan, 2000). During its maturation, the core protein undergoes two consecutive membrane-dependent cleavage events: (i) the first generates the p23 protein, the immature core protein of 191 aa, and is mediated by a signal peptidase; and (ii) the second yields the p21 protein and is mediated by a signal peptide peptidase (McLauchlan *et al.*, 2002). The C terminus of p21 has not been mapped correctly yet and different locations have been reported (Hussy *et al.*, 1996; Liu *et al.*, 1997; Lo *et al.*, 1995; Santolini *et al.*, 1994; Yasui *et al.*, 1998). Both p23 and p21 core proteins have been termed, in some early publications, as p21 and p19, respectively. In addition, a 16 kDa protein, called p16, can be expressed also from the HCV capsid protein-encoding sequence (Lo *et al.*, 1994, 1995; Yeh *et al.*, 2000). The identity of this product remained unclear until the recent observation of a novel translation mechanism

within the capsid-encoding sequence corresponding to a ribosomal frameshift, a mechanism unique among members of the family *Flaviviridae* (Varaklioti *et al.*, 2002; Walewski *et al.*, 2001; Xu *et al.*, 2001). This protein has been called F (frameshifted) (Xu *et al.*, 2001) protein or ARFP (alternative ribosomal frameshift protein) (Walewski *et al.*, 2001).

In this study, we investigated the expression of the F protein in *in vitro* and *in vivo* expression systems in the presence or absence of mutations that have been reported previously to be associated with the expression of the p16 protein (Yeh *et al.*, 2000). In the absence of any mutation, no ribosomal frameshifting was observed. In contrast, a framshifted protein was clearly identified when mutations were introduced at nt 367 and 373 (codons 9 and 11) in the capsid-encoding sequence. Furthermore, the data obtained in time-course experiments revealed that the F protein is a very short-lived protein and that its stability can be maintained by the use of the proteasome inhibitor MG132.

METHODS

Plasmid constructs for frameshifting assays *in vitro* and *in vivo*. Plasmids were constructed using standard methods (Sambrook *et al.*, 1989). HCV sequences were amplified from clones derived from strain H (Feinstone *et al.*, 1981; Fournillier-Jacob *et al.*, 1996). The HCV 5'UTR was introduced into plasmid pTM4/CE_1E_2p7 (Fournillier-Jacob *et al.*, 1996) in place of the encephalomyocarditis virus IRES sequence by PCR with appropriate oligonucleotides and templates. Mutations $G^{367}\rightarrow A$ and $C^{373}\rightarrow A$ were introduced into pTHC/CE_1E_2p7 to generate the plasmid pTHC/CE_1E_2p7m by site-directed mutagenesis using enzymatic inverse PCR, as described by Stemmer & Morris (1992). These plasmids were used as template and also for cloning by replacing the sequence encoding the structural proteins with the new PCR products. Plasmids pTHC/C^{Δ}-Luc^{-} and pTHC/C^{Δ}m-Luc^{-} contain the wild-type and mutant capsid-encoding region (nt 342–585), respectively, from the HCV strain H isolate in which the sequence encoding the firefly luciferase is not fused in-frame with the capsid-encoding sequence or with the F-encoding sequence. The firefly luciferase sequence was amplified by PCR using appropriate primers designed to contain *Asp*718 and *Bgl*II sites at the 5' and 3' ends of the fragment. The restriction site *Bgl*II is an unique restriction site created at the 3' end of the HCV structural sequence. All sequences at the junction between HCV and firefly luciferase sequences are indicated in Fig. 1. To restore a correct reading frame between the sequence encoding the first amino acids of the capsid protein and the firefly luciferase, the plasmids pTHC/C^{Δ}-Luc^{-} and pTHC/C^{Δ}m-Luc^{-} were cleaved at the *Asp*718 restriction site and the sequence was filled with Klenow and then ligated with T4 DNA ligase. The resulting plasmids pTHC/C^{Δ}-Luc^{+} and pTHC/C^{Δ}m-Luc^{+} were obtained. Plasmids pTHC/CF^{Δ}-Luc^{+} and pTHC/CF^{Δ}m-Luc^{+} contain the sequence encoding the firefly luciferase fused in-frame with the sequence encoding the F protein and not the C protein. Cloning was performed by PCR using appropriate oligonucleotides. In all constructs containing the firefly luciferase sequence, the AUG initiation codon of the luciferase gene was omitted during construction in order to avoid all internal translational of this gene in an *in vitro* transcription/translation assay.

Plasmids pTHC/CF-myc and pTHC/CFm-myc contain the entire HCV 5'UTR (nt 1–341) and part of the core-encoding sequence (nt 342–825) from HCV-H genotype 1a fused with the sequence encoding the Myc epitope (EQKLISEEDL) preceded by three Gly residues. This sequence is fused in-phase only with the sequence encoding the HCV F protein. Cloning was performed by PCR using appropriate oligonucleotides and the antisense primer was designed to contain a *Bgl*II

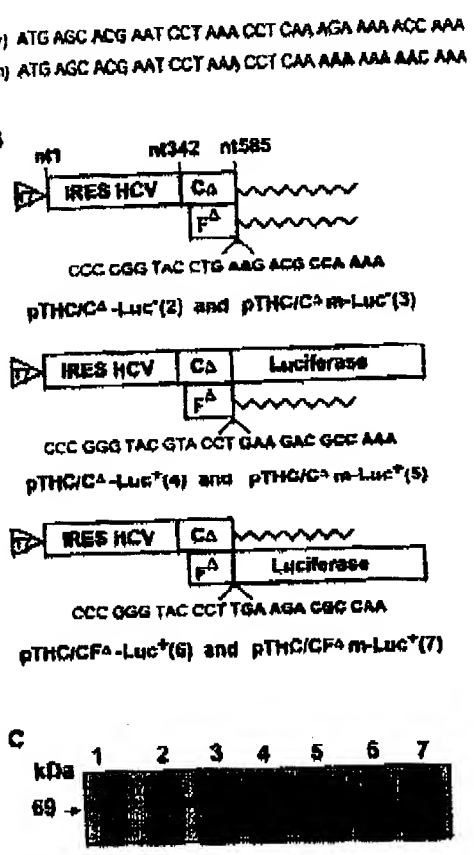

Fig. 1. Tagging experiments with the luciferase gene. (A) The first 12 codons of the capsid protein are shown. Differences between the wild-type (+) and the mutant (m) sequences are indicated. Mutated nucleotides are indicated in bold. (B) Schemanc representation of the constructs used for the tagging experiments with luciferase. The entire HCV IRES and part of the HCV capsid-encoding sequences were cloned into the pTHC plasmid. Nucleotide sequences at the junction of the capsid-encoding region and the luciferase-encoding region are illustrated below. The beginning of the sequence encoding the luciferase gene is shown in bold. Broken lines indicate that the sequence encoding the luciferase protein is out of frame with the sequences encoding either the capsid protein or the F protein. (C) DNA templates were used to programme coupled transcription/translation reactions, as described previously (Meunier *et al.*, 1999). These DNAs correspond, respectively, to pTHC/C^{Δ}-Luc^{-} (lane 2), pTHC/C^{Δ}m-Luc^{-} (lane 3), pTHC/C^{Δ}-Luc^{+} (lane 4), pTHC/C^{Δ}m-Luc^{+} (lane 5), pTHC/CF^{Δ}-Luc^{+} (lane 6) and pTHC/CF^{Δ}m-Luc^{+} (lane 7). The control of the translation assay without RNA is indicated in lane 1. Translation products were analysed by SDS-PAGE (10% polyacrylamide gel). The molecular mass of the fusion protein (in kDa) is indicated on the left.

restriction site and the sequence encoding the Myc epitope. Plasmid pTM1/myc-F contains the entire sequence encoding the F protein (nt 346–825), tagged with the Myc epitope (EQKLISEEDL) at its N terminus. The nucleotide sequence AAAGAAAAA was replaced during PCR by the nucleotide sequence AAGAAGAAG in order to disrupt the A-rich region. Plasmid pTM1/myc-C_{1-163} contains the sequence encoding the first 163 aa and is tagged with the Myc epitope at its N terminus. Each PCR product was cloned into the pTM1 plasmid.

Generation of recombinant vaccinia viruses. Transfection and isolation of recombinant viruses were performed essentially as described (Kieny *et al.*, 1984). The following vaccinia virus recombinant has been described previously: vTF7-3 (expressing the T7 DNA-dependent RNA polymerase) (Fuerst *et al.*, 1986).

Immunoprecipitation and Western blot assays. Cells expressing HCV proteins were metabolically labelled with [^{35}S]protein Labelling mix ($3{\cdot}7 \times 10^6$ Bq ml^{-1}), as described previously (Meunier *et al.*, 1999). Cells were lysed with 0·5% Igepal CA-630 in TBS (50 mM Tris/HCl, pH 7·5, and 150 mM NaCl). Immunoprecipitations were carried out as described (Dubuisson & Rice, 1996). Immune complexes were boiled for 5 min in Laemmli's buffer before analysis by SDS-PAGE. Gels were then treated as described previously (Meunier *et al.*, 1999). Proteins bound to nitrocellulose membranes (PVDF transfer membrane, NEN Life Science Products) were revealed by enhanced chemiluminescence detection (Amersham Pharmacia), as recommended by the manufacturer, with the specific monoclonal antibodies (mAbs) anti-C (diluted 1:5000) (Maillard *et al.*, 2001), anti-myc (diluted 1:200) (ATCC CRL-1725) or anti-luc (diluted 1:1000) (Promega).

Luciferase assays. Luc activity was assayed with the Luciferase Reporter Assay system (Promega) on rabbit reticulocyte lysate expressing the appropriate constructs or on cell lysates provided by cells infected with different recombinant viruses expressing the firefly luciferase activity. HepG2 cells (10^5 cells) plated in a multiwell plate were infected with the appropriate vaccinia virus recombinant at an m.o.i. of 5 p.f.u. per cell. At 6 h post-infection (p.i.), cells were washed twice with PBS, scraped out with 120 µl 1 × Reporter Lysis buffer (Promega), lysed by one freeze–thaw cycle, vortexed and spun at 25 °C to pellet cell debris. Supernatant (4 µl) in 10 µl water was placed in a luminometer (Lumat. LB9501 Berthold) and the reaction was started by injection of 20 µl Luciferase Assay reagent (Promega). Light emission was recorded for 15 s.

RESULTS

In vitro analysis of frameshifting

A recent study on the F protein has indicated that the frameshifting might not be observed in some isolates (Varaldiori *et al.*, 2002). The F protein has indeed been observed clearly in an HCV type 1 isolate but no frameshifting was detected with the HCV strain H isolate. In addition, earlier data have indicated that p16 is produced when mutations are present in the N terminus of the capsid-encoding region (Yeh *et al.*, 2000). These data suggested to us that the sequence present in the N terminus of the capsid-encoding region might influence the efficiency of frameshifting. To test this hypothesis, frameshifting has been analysed in the presence or absence of mutations.

Mutations reported previously – $G^{367}\rightarrow A$ (codon 9) and $C^{373}\rightarrow A$ (codon 11) (Yeh *et al.*, 2000) – were introduced into the N terminus of the capsid-encoding region; this construct will be referred to the mutant 'm' (Fig. 1A). To detect frameshifting and to analyse the level of expression of the framshifted protein, the F protein reading frame was fused in-frame with the Luc-encoding sequence (pTHC/CF^{Δ}-Luc^{+} and pTHC/CF^{Δ}m-Luc^{+} constructs). With this type of approach, the fully active luciferase protein can be detected only if a ribosomal frameshift occurs during translation. To be closer to the context of HCV expression, the sequences encoding the HCV C protein and firefly luciferase were placed under the translational control of the HCV 5′UTR. All DNA constructs used for this study are illustrated in Fig. 1(B). As expected, a band with an apparent molecular mass of 69 kDa and corresponding to the N terminus of the F protein in fusion with Luc (F^{Δ}-Luc) was detected with the *in vitro* transcription/translation assay resulting from the plasmid containing the mutations (pTHC/CF^{Δ}m-Luc^{+}, Fig. 1C). However, no band was observed when no mutation was introduced into the N terminus of the capsid-encoding region (pTHC/CF^{Δ}-Luc^{+}, Fig. 1C). To determine the efficiency of ribosomal frameshifting, the luciferase activity of the F^{Δ}-Luc fusion protein was measured and compared to that of a fusion protein in which the luciferase sequence was fused in-frame with the N-terminal sequence of the C protein (C^{Δ}-Luc). About 30 and 3% of luciferase activity was observed for the mutated and non-mutated F^{Δ}-Luc proteins, respectively (data not shown).

To confirm that the mutations introduced into the N terminus of the capsid-encoding region have no effect on the level of expression of the capsid protein, the Luc-encoding sequence was fused in-frame with the sequence encoding the first 82 aa of the capsid protein (pTHC/C^{Δ}-Luc^{+} and pTHC/C^{Δ}m-Luc^{+}). The fusion product, corresponding to each construct, yielded a 69 kDa protein in an *in vitro* transcription/translation assay (Fig. 1C). Levels of expression were similar and correlated with luciferase activities (data not shown). Additional constructs in which the sequence encoding the first 82 aa of the HCV capsid protein was fused out of frame with the firefly luciferase-encoding sequence (pTHC/C^{Δ}-Luc^{-} and pTHC/C^{Δ}m-Luc^{-}) were used as negative controls. These constructs are not supposed to produce a luciferase fusion protein even if a +1 ribosomal frameshift occurs in the capsid-encoding sequence due to the presence of termination codons in the +1 frame of the luciferase-encoding sequence. This was confirmed by SDS-PAGE analysis of [^{35}S]methionine-labelled translation products obtained in an *in vitro* transcription/translation assay using rabbit reticulocyte lysates (Fig. 1C). Additionally, no firefly luciferase activity associated with these constructs was observed (data not shown).

These results indicate that a frameshifted fusion protein with an active luciferase protein is produced *in vitro* when modifications at nt 367 and 373 (codons 9 and 11) are introduced within the nucleotide sequence encoding the N terminus of the HCV capsid protein. However, in the absence of any mutation, no frameshift is detectable above background level.

In vivo analysis of the frameshifting

Due to the presence of some factors that may influence ribosome–mRNA interactions (Parkin *et al.*, 1992; Reil *et al.*, 1993), it is well known that *in vivo* frameshifting results are generally different from those observed *in vitro*. To confirm the above-mentioned conclusions by *in vivo* experiments, recombinant vaccinia viruses were generated by homologous recombination and used to infect HepG2 cells. The proteins expressed by the different recombinant vaccinia viruses were analysed by SDS-PAGE followed by Western blot analysis using anti-luc or anti-C mAbs (Fig. 2). All constructs analysed *in vitro* were examined *in vivo*. A product fused in-frame with the fully active firefly protein was detected in HepG2 cells infected with the recombinant vaccinia virus vvpTHC/CF$^{\Delta}$m-Luc^{+} (Fig. 2B, lane 7), indicating that frameshifting was observed *in vivo*. In addition, 16 % luciferase activity has been observed for this construct (Fig. 2D, lane 7). In contrast, the firefly luciferase activity was only slightly above background in the absence of any mutation in the F$^{\Delta}$-Luc fusion protein (Fig. 2D, lane 6), which correlates with the absence of detection of the corresponding product by Western blotting (Fig. 2B, lane 6). These data indicate that the mutations introduced in the capsid-encoding sequence are necessary to improve the *in vivo* expression of a frameshifted protein. To confirm that all constructs could be expressed *in vivo*, the presence of capsid-derived proteins was analysed by Western blotting with an anti-C mAb. A truncated form of the capsid protein of the expected size (10 kDa) was detected from cell extracts of HepG2 cells infected with the recombinant vaccinia viruses vvpTHC/C$^{\Delta}$-Luc^{-}, vvpTHC/C$^{\Delta}$m-Luc^{-}, vvpTHC/CF$^{\Delta}$-Luc^{+} and vvpTHC/CF$^{\Delta}$m-Luc^{+} as expected (Fig. 2C, lanes 2, 3, 6 and 7). The small size of these proteins is due to the presence of termination codons in the Luc sequence, which are not in the same frame as the capsid protein. The fused proteins resulting from the capsid protein with the luciferase were only detected when the HepG2 cells were infected with the recombinant viruses vvpTHC/C$^{\Delta}$-Luc^{+} and vvpTHC/C$^{\Delta}$m-Luc^{+} (Fig. 2A, lanes 4 and 5). The corresponding fused proteins were also detected by Western blotting with anti-luc mAb (Fig. 2B, lanes 4 and 5). These constructs gave the highest reproducible levels of luciferase activity (Fig. 2D, lanes 4 and 5).

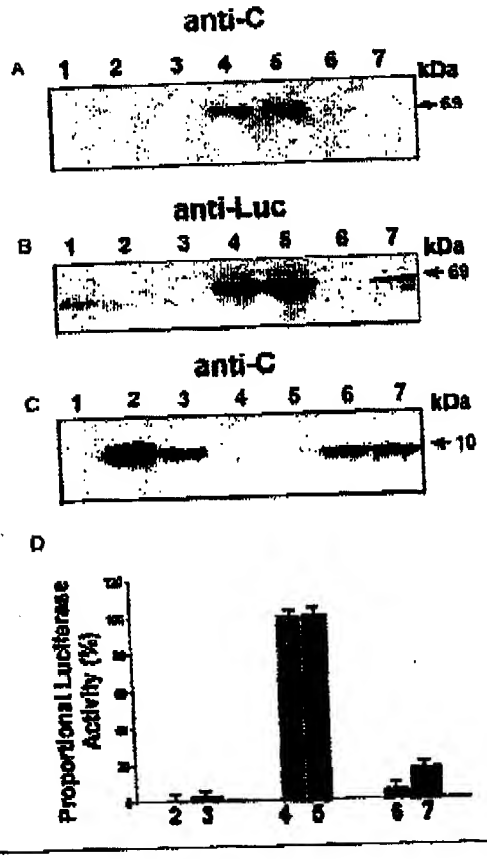

Fig. 2. Detection in HepG2 cells of different products resulting from the tagging experiments with the luciferase gene. HepG2 cells were infected with vTF7-3 alone (lane 1) or co-infected with vTF7-3 and pTHC/C$^{\Delta}$-Luc^{-} (lane 2), vvpTHC/C$^{\Delta}$m-Luc^{-} (lane 3), vvpTHC/C$^{\Delta}$-Luc^{+} (lane 4), vvpTHC/C$^{\Delta}$m-Luc^{+} (lane 5), vvpTHC/CF$^{\Delta}$-Luc^{+} (lane 6) or vvpTHC/CF$^{\Delta}$m-Luc^{+} (lane 7) at an m.o.i. of 5 p.f.u. per cell. At 6 h p.i., cells were lysed in luciferase cell culture lysis reagent. Then, the proteins of the cell lysates were separated by SDS-PAGE (10% polyacrylamide gel) and revealed by Western blotting using anti-C mAb (A, C) and anti-luc mAb (B). Of the cell extracts, 4 µl was measured for luciferase activity, according to the manufacturer's protocol (Promega), and the resulting values in percentages are illustrated in (D). The molecular masses (kDa) of the different proteins are indicated on the right.

Detection of the F protein in HepG2 cells

Taking advantage of the data described above, and to characterize further the expression of the F protein in mammalian cells, plasmids containing the sequence encoding the first 163 aa of the capsid protein downstream of the HCV 5'UTR were constructed with or without mutations in the sequence of the HCV capsid protein. Because no antibodies were available against the F protein, the sequence of the Myc epitope was introduced in-frame with the C-terminal sequence of the predicted F protein (Xu *et al.*, 2001). Plasmids with wild-type (+) and mutant (m) sequences, corresponding to pTHC/CF-myc and pTHC/CFm-myc, respectively, were constructed (Fig. 3A). To analyse the expression of the F protein in mammalian cells, HepG2 cells were infected with recombinant vaccinia viruses vvpTHC/CF-myc and vvpTHC/CFm-myc. The products expressed by these viruses were analysed by SDS-PAGE followed by Western blotting using anti-C or anti-myc mAbs. As shown in Fig. 3(C, lane 3), a 17 kDa product corresponding to the size of the F protein was revealed with the anti-myc mAb in cells infected with vvpTHC/CFm-myc. In contrast, no product was detected with this mAb when the cells were infected with vvpTHC/CF-myc (Fig. 3C, lane

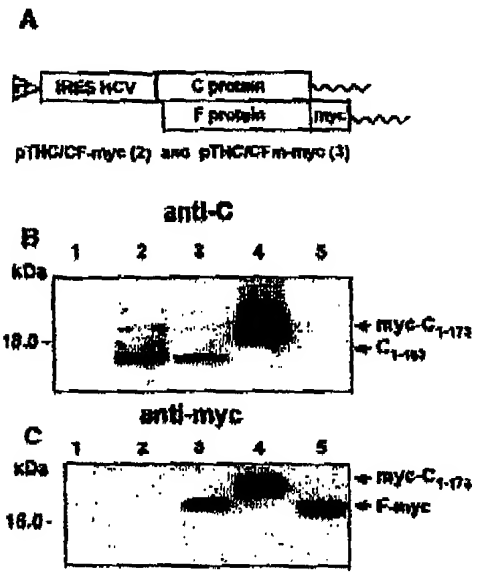

Fig. 3. Tagging experiments with the Myc epitope. (A) Schematic representation of the constructs used for the tagging experiments with *myc*. The entire HCV IRES and part of the HCV capsid-encoding sequences were cloned into pTHC. The C terminus of the F protein was tagged with the Myc epitope. (B, C) HepG2 cells were infected with vTF7-3 alone (lane 1) or co-infected vTF7-3 and vvpTHC/CF-*myc* (lane 2), vvpTHC/CFm-myc (lane 3), vvmyc-C_{1-173} (lane 4) or vvmyc-F_{1-161} (lane 5) at an m.o.i. of 5 p.f.u. per cell. Infected cells were lysed at 6 h p.i. Proteins were separated by SDS-PAGE and then revealed by Western blotting using anti-C *mAb* (B) and anti-*myc* mAb (C).

2). In addition, a band of 18 kDa corresponding to a truncated form of the capsid protein (C_{1-163}) was revealed with the anti-C mAb (Fig. 3B, lanes 2 and 3). Recombinant vaccinia viruses expressing HCV C_{1-173} or F_{1-161} proteins tagged at their N terminus with the Myc epitope were used as controls. The sequence encoding the F_{1-161} protein was based on its reading frame without the first amino acids of the HCV capsid protein. Modifications in the stretch of A residues were introduced in the sequence in order to avoid a possible ribosomal frameshift (codons 9–14). As shown in Fig. 3(B, C, lanes 5), the F_{1-161} protein (17 kDa) was detected only by anti-myc mAb, while the tagged capsid protein was detected with the anti-C and anti-myc mAbs (Fig. 3B, C, lane 4).

These data are in agreement with the *in vitro* experiments and confirm that a frameshift occurs *in vivo* at least when some mutations are introduced in the capsid-encoding sequence. In this particular case, modifications of G→A and C→A at nt 367 and 373, respectively, lead to the formation of a perfect heptanucleotide sequence, which could be considered as the first element necessary for the ribosomal frameshift and could be assimilated to the slippery sequence (see Discussion).

Analysis of F protein stability

We also wanted to determine whether the F protein is stable during its translation in mammalian cells. The expression of the F protein was analysed in pulse–chase experiments by infecting HepG2 cells with the recombinant vaccinia virus vvpTHC/CFm-myc. As shown in Fig. 4(A), the level of expression of the F protein decreased very quickly. After 15 min of chase, the intensity of the band was already very low. The estimated half-life of the F protein was approximately 10 min (Fig. 4A). A decrease in the amount of the F protein was also observed when the HepG2 cells were infected with the recombinant virus vvmyc-F_{1-161} (Fig. 4B). However, after 2 h of chase, 20 % of the F protein was still detected, suggesting that overexpression of the F protein might reduce its degradation. The capsid protein tagged with the Myc epitope at its N terminus (myc-C_{1-173}) was used as a control. As shown in Fig. 4(C), this protein was very stable during the same pulse–chase conditions.

Degradation of most proteins in mammalian cells occurs via the proteasome (Lee & Goldberg, 1998), which degrades both short-lived proteins ($t_{1/2}$ < 3 h) and more stable proteins ($t_{1/2}$ of h or days). Investigation of protein turnover has been facilitated by specific proteasome inhibitors like lactacystin or MG132. In this study, we wanted to address

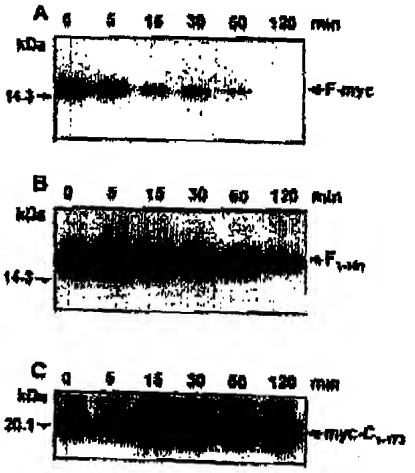

Fig. 4. Expression of F-myc, myc-C_{1-173} and myc-F_{1-161} analysed in pulse–chase experiments. HepG2 cells were co-infected with vTF7-3 and (A) vvpTHC/CFm-myc, (B) vvmyc-F_{1-161} or (C) vvmyc-C_{1-173} at an m.o.i. of 5 p.f.u. per cell. Infected cells were pulse-labelled for 10 min and chased for the times indicated. Cell lysates were immunoprecipitated with anti-*myc* mAb. Samples were separated by SDS-PAGE (15% polyacrylamide gel).

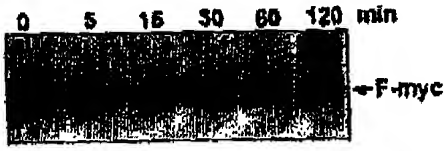

Fig. 5. Pulse–chase analysis of the F-myc protein with the proteasome inhibitor MG132. HepG2 cells were co-infected with vTF7-3 and vvpTHC/CFm-myc at an m.o.i. of 5 p.f.u. per cell. At 1 h before pulse labelling, MG132 (2 μM) was added to the infected cells. Infected cells were then pulse-labelled for 10 min and chased, as indicated before, in the presence of inhibitor. Cell lysates were immunoprecipitated with anti-myc mAb. Samples were separated by SDS-PAGE (15% polyacrylamide gel). A control experiment in the absence of MG132 was also carried out (data not shown).

the question of whether the proteasome pathway could be involved in the degradation of this protein. The proteasome inhibitor MG132 was used during infection of HepG2 cells with both recombinant vaccinia viruses vTF7-3 and vvpTHC/CFm-myc and cells were pulse-labelled as described above. As shown in Fig. 5, the degradation of the F protein was markedly inhibited by adding the proteasome inhibitor. Indeed, the amount of F protein was stable during the pulse–chase experiment. Thus, these results seem to indicate a role of the proteasome pathway in the turnover of this protein. Based on these results, we reinvestigated the expression of a frameshifted protein in the absence of any mutation in the presence of the proteasome inhibitor MG132. Indeed, it is possible that the absence of detection of the F protein with the wild-type construct could be due to a rapid degradation of this protein during its expression. The pulse–chase experiment was done under the same conditions as described above. However, even in the presence of the proteasome inhibitor, the F protein was not detected (data not shown). These data indicate that the lack of detection of the F protein in the absence of any mutation is not due to a rapid degradation of this protein.

Intracellular localization of F protein

As a next step in the characterization of the F protein, we analysed its subcellular localization. Fig. 6 illustrates the immunostaining of vvpTHC/CF-myc (Fig. 6C, D) and vvpTHC/CFm-myc (Fig. 6E, F) infected HepG2 cells observed by immunofluorescence using the anti-C mAb (Fig. 6A, C and E) or anti-myc mAb (Fig. 6B, D and F). As expected, immunofluorescence staining with anti-myc mAb was detected with the vvpTHC/CFm-myc recombinant virus (Fig. 6F). Reactivity was localized mainly in the cytoplasm and in some cells the reactivity had a perinuclear distribution. The immunostaining experiment with the anti-C mAb was used as a control of the expression of the capsid protein. Reactivity against the capsid protein was localized in the cytoplasm. Reactivity of the anti-myc and anti-C mAbs was also performed with control vTF7-3-infected cells in order to evaluate the level of background immunostaining (Fig. 6A, B).

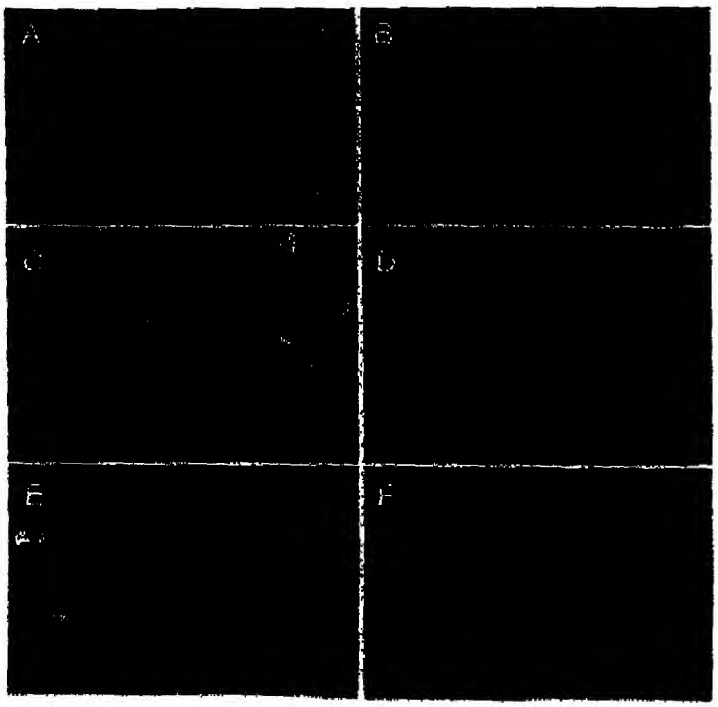

Fig. 6. Detection of the F protein by indirect immunofluorescence. Subconfluent HepG2 cells grown on coverslips were infected with vTF7-3 alone (A, B), vTF7-3 and vvpTHC/CF-myc (C, D) or vTF7-3 and vvpTHC/CFm-myc (E, F) at a m.o.i. of 5 p.f.u. per cell. At 8 h p.i., cells were fixed for 10 min with paraformaldehyde (4% in PBS), permeabilized for 30 min with Triton X-100, and labelled with anti-C mAb (diluted 1:300) (secondary antibody: Rhodamine Red-X-conjugated $F(ab')_2$ fragment donkey anti-mouse IgG (A, C and E) or anti-myc mAb (diluted 1:100) (B, D and F)).

DISCUSSION

Our data reveal that the F protein, produced by a ribosomal frameshift mechanism, is expressed in mammalian cells. Translation of this protein is favoured by mutations located in the capsid-encoding sequence. When expressed alone, this protein is short-lived compared to the other HCV proteins (Pietschmann *et al.*, 2001) and its degradation involves the proteasome pathway.

Reporter enzymes, such as luciferase, CAT and β-galactosidase, are used commonly *in vivo* and *in vitro* (Naylor, 1999). Such enzymes are not expressed naturally in mammalian cells, making them useful to study gene expression in cells. Luciferase-tagging experiments used in this study have revealed that the ribosomal frameshift is more efficient when mutations at nt 367 (codon 9) and 373 (codon 11) are introduced in the capsid-encoding sequence compared to the results obtained with the wild-type sequence. The luciferase activity determined *in vivo* was about 16% relative to the control construct and was less than that observed *in vitro* (30%). This lower efficiency observed in mammalian cells is not unusual and can depend on the genes and the expression systems used (Ivanov *et al.*, 2000). In our study, tagging the C terminus of the F protein has made its detection easier in mammalian cells by immunoblot or immunoprecipitation analyses. Our results have shown that, besides its expression in an *in vitro* system, the F protein can be expressed also in mammalian cells. These data are reinforced by the fact that antibodies directed against the F protein are present in some patient sera, indicating that the F protein is produced during a natural HCV infection in patients (Varaklioti *et al.*, 2002; Xu *et al.*, 2001).

The molecular mechanism of frameshifting leading to the translation of the HCV F protein remains to be determined. However, given the data presented in this study, it could be concluded that the modifications at nt 367 and 373 do not simply increase efficiency but are necessary for the production of the F protein. Interestingly, a similar observation was made using the sequence of a HCV genotype 1b (data not shown). Modifications of codons Arg→Lys (codon 9) and Thr→Asn (codon 11) lead to the modification of a G→A (nt 367) and a C→A (nt 373), generating an A-rich region (10 A residues) between nt 363 and 374 of the HCV sequence. In this context, a slippery sequence can emerge from this region. Indeed, defined initially by *in vitro* translation assays, two structural motifs in mRNA have been characterized as important for an efficient −1 ribosomal frameshift (Brierley, 1995; Dinman, 1995; Gesteland & Atkins, 1996). One is the slippery sequence, the heptanucleotide XXXYYYZ (X is any base, Y is A or U and Z is not G) (Jacks *et al.*, 1988), and the other component is a downstream RNA structural element, either a simple hairpin structure or, more frequently, a pseudoknot (Chamorro *et al.*, 1992; Jacks *et al.*, 1988; ten Dam *et al.*, 1990). In contrast, the +1 ribosomal frameshift in the yeast retrotransposon Ty requires only a short slippery sequence with a rare codon in the original reading frame (Belcourt & Farabaugh, 1990). Frameshifting in human immunodeficiency virus requires only the short slippery sequence but not the 3′ sequence with its predicted stem–loop structure (Wilson *et al.*, 1988). This unusual process occurs also among other viruses, including coronaviruses (Brierley, 1995; Herold & Siddell, 1993) and human astrovirus serotype 1 (Marczinke *et al.*, 1994). In HCV, a single mutation of the codon AGA (Arg) to AAA (Lys) is also sufficient to generate *in vitro* the expression of the F protein (Lo *et al.*, 1994). This mutation creates a heptanucleotide sequence between nt 364 and 372 or 368 and 374. As mutations introduced in this region may recreate a functional slippery sequence, as described for other viruses, it can be suggested that this sequence constitutes the first control element for the translation of the F protein. Comparative studies of the sequence of the codons 8–14 reveal that these codons are also very conserved among the HCV sequences, reinforcing the importance of this region (Rijnbrand & Lemon, 2000). However, if the slippery sequence can be suggested, no available data allow us to determine whether an RNA structural element is also necessary for the ribosomal frameshift.

Interestingly, the F protein is very unstable when expressed in mammalian cells. The degradation of most proteins in mammalian cells occurs via the ubiquitin-proteasome pathway (Lee & Goldberg, 1998). Under these circumstances, ubiquitin is linked covalently to the proteins and is then targeted to a large multiproteinase complex (700 kDa) that constitutes the catalytic core of the 26S proteasome used as an intracellular protein-degrading machine of eukaryotic organisms. The proteasome is essential for the normal turnover of regulatory proteins that controls cell growth and metabolism or is necessary for the removal of damaged or mutated proteins (Molinari *et al.*, 1999). Since the discovery that lactacystin is a potent inhibitor of the 26S proteasome (Fenteany *et al.*, 1995), different novel proteasome inhibitors, like MG132 or proteasome inhibitor I, were developed. In our study, pulse–chase assays indicate that the decreased protein level of the HCV F protein in mammalian cells is due to proteasome degradation. A specific inhibitor of 26S proteasome activity blocks degradation of the F protein during its translation and stabilizes the level of its expression. It is possible that the F protein expressed under our experimental conditions is not folded properly and is, therefore, targeted directly for degradation, as observed for some damaged or mutated proteins (Molinari *et al.*, 1999). In this way, it would be very interesting also to determine whether HCV proteins could be involved in the stabilization of the F protein. Alternatively, it is not the F protein itself but some of its degradation product(s) that might play a functional role in the HCV life cycle.

No function has been attributed to the F protein yet. Its localization in the cytosol suggests that this protein plays a functional role in this compartment. However, we cannot exclude that it moves to another compartment, e.g. the nucleus, after interacting with a cellular component. Its very

low level of expression in the absence of any mutation suggests that the F protein might also be expressed at very low levels in HCV-infected cells, which is not in favour for a role of this protein in virus assembly. Interestingly, mutations leading to higher levels of expression of the F protein *ex vivo* can be observed in HCV isolated from chronically infected patients with HCV-related hepatocellular carcinoma, leading to the hypothesis that higher levels of F protein expression might be linked to virus pathogenesis. Further studies on the F protein should lead to a better understanding of the role of this protein in the HCV life cycle and in pathogenesis.

ACKNOWLEDGEMENTS

We would like to thank Dr Delagneau for the generous gift of anti-C monoclonal antibody. This work was supported by the CNRS (Centre National de la Recherche Scientifique) and the Institut Pasteur of Lille. Juliette Roussel is recipient of funding from the ANRS (Agence Nationale de la Recherche sur le SIDA).

REFERENCES

Belcourt, M. F. & Farabaugh, P. J. (1990). Ribosomal frameshifting in the yeast retrotransposon Ty: tRNAs induce slippage on a 7 nucleotide minimal site. *Cell* 62, 339–352.

Brierley, I. (1995). Ribosomal frameshifting viral RNAs. *J Gen Virol* 76, 1885–1892.

Chamorro, M., Parkin, N. & Varmus, H. E. (1992). An RNA pseudoknot and an optimal heptameric shift site are required for highly efficient ribosomal frameshifting on a retroviral messenger RNA. *Proc Natl Acad Sci U S A* 89, 713–717.

Choo, Q.-L., Kuo, G., Weiner, A. J., Overby, L. R., Bradley, D. W. & Houghton, M. (1989). Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. *Science* 244, 359–362.

Choo, Q.-L., Richman, K. H., Han, J. H. & 11 other authors (1991). Genetic organization and diversity of the hepatitis C virus. *Proc Natl Acad Sci U S A* 88, 2451–2455.

Dinman, J. D. (1995). Ribosomal frameshifting in yeast viruses. *Yeast* 11, 1115–1127.

Dubuisson, J. & Rice, C. M. (1996). Hepatitis C virus glycoprotein folding: disulfide bond formation and association with calnexin. *J Virol* 70, 778–786.

Feinstone, S. M., Alter, H. J., Dienes, H. P., Shimizu, Y., Popper, H., Blackmore, D., Sly, D., London, W. T. & Purcell, R. H. (1981). Non-A, non-B hepatitis in chimpanzees and marmosets. *J Infect Dis* 144, 588–598.

Fenteany, G., Standaert, R. F., Lane, W. S., Choi, S., Corey, E. J. & Schreiber, S. L. (1995). Inhibition of proteasome activities and subunit-specific amino-terminal threonine modification by lactacystin. *Science* 268, 726–731.

Fournillier-Jacob, A., Cahour, A., Escriou, N., Girard, M. & Wychowski, C. (1996). Processing of the E1 glycoprotein of hepatitis C virus expressed in mammalian cells. *J Gen Virol* 77, 1055–1064.

Fuerst, T. R., Niles, E. G., Studier, F. W. & Moss, B. (1986). Eukaryotic transient-expression system based on recombinant vaccinia virus that synthesizes bacteriophage T7 RNA polymerase. *Proc Natl Acad Sci U S A* 83, 8122–8126.

Gesteland, R. F. & Atkins, J. F. (1996). Recoding: dynamic reprogramming of translation. *Annu Rev Biochem* 65, 741–768.

Herold, J. & Siddell, S. G. (1993). An 'elaborated' pseudoknot is required for high frequency frameshifting during translation of HCV 229E polymerase mRNA. *Nucleic Acids Res* 21, 5838–5842.

Hussy, P., Langen, H., Mous, J. & Jacobsen, H. (1996). Hepatitis C virus core protein: carboxy-terminal boundaries of two processed species suggest cleavage by a signal peptide peptidase. *Virology* 224, 93–104.

Ivanov, I. P., Matsufuji, S., Murakami, Y., Gesteland, R. F. & Atkins, J. F. (2000). Conservation of polyamine regulation by translational frameshifting from yeast to mammals. *EMBO J* 19, 1907–1917.

Jacks, T., Madhani, H. D., Masiarz, F. R. & Varmus, H. E. (1988). Signals for ribosomal frameshifting in the Rous sarcoma virus *gag–pol* region. *Cell* 55, 447–458.

Kieny, M.-P., Lathe, R., Drillien, R., Spehner, D., Skory, S., Schmitt, D., Wiktor, T., Koprowski, H. & Lecocq, J.-P. (1984). Expression of rabies virus glycoprotein from a recombinant vaccinia virus. *Nature* 312, 163–166.

Lee, D. H. & Goldberg, A. L. (1998). Proteasome inhibitors: valuable new tools for cell biologists. *Trends Cell Biol* 8, 397–403.

Lindenbach, B. D. & Rice, C. M. (2001). *Flaviviridae*: the viruses and their replication. In *Fields Virology*, 4th edn, pp. 991–1042. Edited by D. M. Knipe & P. M. Howley. Philadelphia: Lippincott Williams & Wilkins.

Liu, Q., Tackney, C., Bhat, R. A., Prince, A. M. & Zhang, P. (1997). Regulated processing of hepatitis C virus core protein is linked to subcellular localization. *J Virol* 71, 657–662.

Lo, S. Y., Selby, M., Tong, M. & Ou, J. H. (1994). Comparative studies of the core gene products of two different hepatitis C virus isolates: two alternative forms determined by a single amino acid substitution. *Virology* 199, 124–131.

Lo, S. Y., Masiarz, F., Hwang, S. B., Lai, M. M. & Ou, J. H. (1995). Differential subcellular localization of hepatitis C virus core gene products. *Virology* 213, 455–461.

Lohmann, V., Korner, F., Koch, J., Herian, U., Theilmann, L. & Bartenschlager, R. (1999). Replication of subgenomic hepatitis C virus RNAs in a hepatoma cell line. *Science* 285, 110–113.

Maillard, P., Krawczynski, K., Nitkiewicz, J. & 7 other authors (2001). Nonenveloped nucleocapsids of hepatitis C virus in the serum of infected patients. *J Virol* 75, 8240–8250.

Marczinke, B., Bloys, A. J., Brown, T. D., Willcocks, M. M., Carter, M. J. & Brierley, I. (1994). The human astrovirus RNA-dependent RNA polymerase coding region is expressed by ribosomal frameshifting. *J Virol* 68, 5588–5595.

McHutchison, J. G. & Poynard, T. (1999). Combination therapy with interferon plus ribavirin for the initial treatment of chronic hepatitis C. *Semin Liver Dis* 19, 57–65.

McHutchison, J. G., Gordon, S. C., Schiff, E. R. & 7 other authors (1998). Interferon α2b alone or in combination with ribavirin as initial treatment for chronic hepatitis C. Hepatitis Interventional Therapy Group. *N Engl J Med* 339, 1485–1492.

McLauchlan, J. (2000). Properties of the hepatitis C virus core protein: a structural protein that modulates cellular processes. *J Viral Hep* 7, 2–14.

McLauchlan, J., Lemberg, M. K., Hope, G. & Martoglio, B. (2002). Intramembrane proteolysis promotes trafficking of hepatitis C virus core protein to lipid droplets. *EMBO J* 21, 3980–3988.

Meunier, J. C., Fournillier, A., Choukhi, A., Cahour, A., Cocquerel, L., Dubuisson, J. & Wychowski, C. (1999). Analysis of the glycosylation sites of hepatitis C virus (HCV) glycoprotein E1 and the influence of

E1 glycans on the formation of the HCV glycoprotein complex. *J Gen Virol* 80, 887–896.

Miller, R. H. & Purcell, R. H. (1990). Hepatitis C virus shares amino acid sequence similarity with pestiviruses and flaviviruses as well as members of two plant virus supergroups. *Proc Natl Acad Sci U S A* 87, 2057–2061.

Molinari, E., Gilman, M. & Natesan, S. (1999). Proteasome-mediated degradation of transcriptional activators correlates with activation domain potency *in vivo*. *EMBO J* 18, 6439–6447.

Naylor, L. H. (1999). Reporter gene technology: the future looks bright. *Biochem Pharmacol* 58, 749–757.

Parkin, N. T., Chamorro, M. & Varmus, H. E. (1992). Human immunodeficiency virus type 1 *gag–pol* frameshifting is dependent on downstream mRNA secondary structure: demonstration by expression *in vivo*. *J Virol* 66, 5147–5151.

Pietschmann, T., Lohmann, V., Rutter, G., Kurpanek, K. & Bartenschlager, R. (2001). Characterization of cell lines carrying self-replicating hepatitis C virus RNAs. *J Virol* 75, 1252–1264.

Reil, H., Kollmus, H., Weidle, U. H. & Hauser, H. (1993). A heptanucleotide sequence mediates ribosomal frameshifting in mammalian cells. *J Virol* 67, 5579–5584.

Rijnbrand, R. C. & Lemon, S. M. (2000). Internal ribosome entry site-mediated translation in hepatitis C virus replication. *Curr Top Microbiol Immunol* 242, 85–116.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: a Laboratory Manual*, 2nd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Santolini, E., Migliaccio, G. & La Monica, N. (1994). Biosynthesis and biochemical properties of the hepatitis C virus core protein. *J Virol* 68, 3631–3641.

Schalm, S. W., Weiland, O., Hansen, B. E. & 9 other authors (1999). Interferon-ribavirin for chronic hepatitis C with and without cirrhosis: analysis of individual patient data of six controlled trials. Eurohep Study Group for Viral Hepatitis. *Gastroenterology* 117, 408–413.

Stemmer, W. P. C. & Morris, S. K. (1992). Enzymatic inverse PCR: a restriction site independent, single-fragment method for high-efficiency, site directed mutagenesis. *Biotechniques* 13, 214–220.

Takeuchi, K., Kubo, Y., Boonmar, S. & 7 other authors (1999). The putative nucleocapsid and envelope protein genes of hepatitis C virus determined by comparison of the nucleotide sequences of two isolates derived from an experimentally infected chimpanzee and healthy human carriers. *J Gen Virol* 71, 3027–3033.

ten Dam, E. B., Pleij, C. W. & Bosch, L. (1990). RNA pseudoknots: translational frameshifting and readthrough on viral RNAs. *Virus Genes* 4, 121–136.

Varaklioti, A., Vassilaki, N., Georgopoulou, U. & Mavromara, P. (2002). Alternate translation occurs within the core coding region of the hepatitis C viral genome. *J Biol Chem* 277, 17713–17721.

Walewski, J. L., Keller, T. R., Stump, D. D. & Branch, A. D. (2001). Evidence for a new hepatitis C virus antigen encoded in an overlapping reading frame. *RNA* 7, 710–721.

Wilson, W., Braddock, M., Adams, S. E., Rathjen, P. D., Kingsman, S. M. & Kingsman, A. J. (1986). HIV expression strategies: ribosomal frameshifting is directed by a short sequence in both mammalian and yeast systems. *Cell* 55, 1159–1169.

Xu, Z., Choi, J., Yen, T. S., Lu, W., Strohecker, A., Govindarajan, S., Chien, D., Selby, M. J. & Ou, J. (2001). Synthesis of a novel hepatitis C virus protein by ribosomal frameshift. *EMBO J* 20, 3840–3848.

Yasui, K., Wakita, T., Tsukiyama-Kohara, K., Funahashi, S. I., Ichikawa, M., Kajita, T., Moradpour, D., Wands, J. R. & Kohara, M. (1998). The native form and maturation process of hepatitis C virus core protein. *J Virol* 72, 6048–6055.

Yeh, C. T., Lo, S. Y., Dai, D. L., Tang, J. H., Chu, C. M. & Liaw, Y. F. (2000). Amino acid substitutions in codons 9–11 of hepatitis C virus core protein lead to the synthesis of a short core protein product. *J Gastroenterol Hepatol* 15, 182–191.

Appendix C
USSN 09/719,277

Volume 46(5): 653–660, 1998
The Journal of Histochemistry & Cytochemistry
http://www.jhc.org

ARTICLE

Detection and Localization by In Situ Molecular Biology Techniques and Immunohistochemistry of Hepatitis C Virus in Livers of Chronically Infected Patients

Francine M. Walker, Marie-Christine Dazza, Marie-Christine Dauge, Olivier Boucher, Christophe Bedel, Dominique Henin, and Therese Lehy

Department of Pathology (FMW,MCD,OB,CB,DH), INSERM U10 (TL), and INSERM U13 (MCD), Hopital Bichat-Claude Bernard, Paris, France

SUMMARY Hepatitis C virus (HCV) detection in the livers of chronically infected patients remains a debatable issue. We used immunohistochemistry, in situ hybridization (ISH) alone or after microwave heating with FITC-labeled probes, RT-PCR with unlabeled primers followed by ISH (RT-PCR–ISH), and in situ RT-PCR with FITC-labeled primers (in situ RT-PCRd) to localize the virus in 38 liver biopsy specimens from 21 chronically infected HCV patients treated with interferon-α (IFN-α). Biopsies were taken at the beginning and end of IFN-α treatment and 1 year later. Results were compared with that of HCV-PCR in serum. RT-PCR–ISH and in situ RT-PCRd showed HCV signal in all liver biopsies even in responders with seronegative HCV PCR. This signal was intranuclear, diffuse, or peripheral, in hepatocytes, bile ductule cells, and lymphocytes. Cytoplasmic signals were occasionally observed. Whereas the percentage of labeled hepatocytes remained constant, the number of labeled lymphoid follicles decreased after INF-α therapy. Immunohistochemistry resulted in the same pattern of positivity but it was weaker and inconstant. This study indicates the persistency of HCV latency in IFN-α responders 1 year after IFN-α treatment cessation, a finding that certainly deserves confirmation. (J Histochem Cytochem 46:653–660, 1998)

KEY WORDS
HCV
interferon-α therapy
in situ RT-PCR
in situ RT-PCRd
RT-PCR–ISH
immunohistochemistry
C100-3 protein
human liver

HEPATITIS C VIRUS (HCV), a positive-stranded RNA virus and member of the *Flaviridae* family, has been reported to be the major causative agent of chronic HCV hepatitis. Liver-related death has been reported to occur in 20–25% of patients with cirrhosis. Interferon-α (IFN-α) is currently used as an antiviral agent for the treatment of chronic hepatitis C. However, only 50% of the treated patients respond to therapy and only 20% sustain a response (Houghton 1996). Although several techniques have been used to localize HCV in the liver, their reliability has not yet been proved. Immunohistochemistry (IHC) has limited sensitivity for detecting HCV proteins. In situ hybridization (ISH) is also insensitive because HCV replication is very low, estimated at about 20 copies per cell, which is at the detection limit of nonradioactive probes (Weiner et al. 1990). Recently, in situ amplification techniques after reverse transcriptase (in situ RT-PCR) were tested either with direct incorporation of labeled nucleotides or primers (in situ RT-PCRd) or with unlabeled primers and nucleotides followed by ISH (RT-PCR–ISH). Intracellular viral localization by these different techniques has led to conflicting results: it may be peri- or intranuclear (Nuovo 1994) or cytoplasmic (Barba et al. 1997). HCV has been detected in hepatocytes alone or in some bile duct cells, Kupffer cells, and portal lymphocytes.

The aims of the present study were to localize HCV by in situ molecular biology techniques (ISH, in situ RT-PCRd, and RT-PCR–ISH) in the livers of patients with chronic hepatitis C, to compare results obtained by these techniques with IHC detection of the viral protein (the superoxide dismutase–HCV recombinant polypeptide C100-3), and to compare these data to

Correspondence to: Dr. Francine M. Walker, Service d'Anatomie Pathologique, Hopital Bichat-Claude Bernard, 46 rue Henri Huchard, 75877 Paris Cedex 18, France

Received for publication June 2, 1997; accepted November 21, 1997 (7A4354)

0022-1554/98/$3.30

those of RT-PCR detection of HCV in serum before and after IFN-α therapy.

Materials and Methods

Patient Selection

Twenty-one white HCV-confirmed seropositive patients (11 women and 10 men) were selected for this study, as all had sera collected at the time of each liver biopsy. The median age was 33 years (range 21–60 years). Patients received the same therapeutic treatment (INF-α) for 6 or 12 months. The criterion for INF-α treatment was a persistently elevated serum alanine transaminase (ALT) for at least 6 months. None of the patients was seropositive for HIV or HBV. None had received previous treatment. For each patient, ALT level determination was performed at the same time as biopsy.

Tissue Sample Processing

Liver biopsies were attempted at three time points: just before INF-α treatment (B1) ($n = 8$), at treatment termination (B2) ($n = 13$), and 1 year after treatment termination (B3) ($n = 17$). In total, 38 biopsy specimens were studied. Five patients were able to tolerate all three biopsies, seven underwent initial and second biopsies, and 10 underwent the second and third biopsies.

Tissue specimens were fixed in 10% formalin and then paraffin-embedded. Liver samples from alcoholic and hepatitis B patients seronegative for HCV antibodies and for HCV RNA, as well as heterologous tissue samples (e.g., salivary glands and gastric tissue) from these same patients, were used as negative controls. Intrahepatic HCV detection was compared with HCV detection in each corresponding serum sample.

Serum Evaluation of HCV

Two methods were used on frozen serum specimens. (a) HCV RNA quantification by the bDNA method, only positive for at least 0.35×10^6 viral copies/ml, was performed with the quantitative bDNA signal amplification assay (Quantiplex TM HCV RNA 1.0; Chiron Diagnostics, Emeryville, CA) at each time point. All samples were run in duplicate. (b) If the bDNA quantification remained negative, the Nested RT-PCR was tested (Amplicor TM HCV; Roche, Paris, France).

In Situ Molecular Biology Techniques. Primers and probes were all located in the same 5′ non-coding region of the HCV genome. Unlabeled and 5′ FITC-labeled oligonucleotide primers were used. First, 20-mer primers were synthetized by Genset (Paris, France). They trigger amplification of a 260-BP product and correspond to the outer oligonucleotides sequences used by Feray et al. (1992): sense, 5′ GCC ATG GCG TTA GTA TGA GT 3′ (recognizing codons 260–241) and anti-sense, 5′ TGC ACG GTC TAC GAG ACC TC 3′ (recognizing codons 22 to 3). The 36-mers ISH 5′ FITC-labeled probe complementary to the amplicon was 5′ ACA CCG GAA TTG CCA GGA CGA CCG CGT CCT TTC TTG 3′ (codons 177–142). Other commercially produced oligonucleotide primers (either 5′ FITC-labeled or unlabeled) were provided by Argene (Varilles, France): sense (MJIC, 21-mer), 5′ GAC GAA CTA CTG TCT TCA CGC 3′ codons 292–271) and anti-sense (MJ2D, 22-mer), TGC ACG GTC TAC GAG ACC TCC C (codons 24 to 2). In this case, two different 5′ FITC-labeled probes KB2 and KB5, 18 mers each, were used, complementary to the latter amplicon. As internal positive control we used the β-actin RNA.

In situ hybridization was performed in a moist chamber for 18 hr at 37C with all the FITC-labeled probes. About 20 ng of probe was applied per section. To detect the hybridization complex, slides were incubated first with mouse anti-FITC antiserum, second with rabbit anti-mouse immunoglobulins, and then with a mouse anti-rabbit monoclonal antibody linked to APAAP complex (Dakopatts, Glostrup, Denmark). Alkaline phosphatase activity was revealed using NBT-BCIP chromogenic substrate. In some cases, citrate antigen retrieval by microwave heating before ISH was performed.

In Situ RT-PCR Method. As recommended by Nuovo (1994), tissue sections were treated overnight with an RNase-free DNase solution (1 U/ml at 37C; Boehringer Mannheim, Meylan, France) before reverse transcription. They were then incubated in a thermocycler (PTC 100, Research Inc, Watertown, MA) with 10 ml of a solution containing buffer, $MgCl_2$, the anti-sense primer (1 mM), reverse transcriptase, RNase inhibitor, and dNTPs (stock solution 10 mM each) according to the manufacturer's recommendation (RT-PCR kit provided by Perkin–Elmer, Norwalk, CT). PCR amplification was performed in all samples according to our previously published protocol (Walker et al. 1996) with unlabeled or labeled primers. For most samples, amplification was also performed using another thermocycler (Perkin–Elmer). In situ RT-PCR products with labeled primers were directly detected, and RT-PCR products with an unlabeled set of primers were revealed after ISH.

As controls for in situ PCR, liver tissues removed from HCV-positive patients who had a high expression of viral nucleic acids served as initial positive controls. Negative controls included (a) omission of the reverse transcription step and/or DNase pretreatment, (b) incubation of slides in an RNase solution immediately before reverse transcription, (c) for the amplification, omission of the DNA polymerase or primers and incubation of the liver tissue with DNase before PCR amplification and ISH, and (d) the use on HCV infected liver of heterologous human papillomavirus-related primers for amplification followed by ISH with HCV probes or of heterologous HBV probe after amplification with HCV sets of primers. HCV primers and probes were also applied to liver and heterologous tissues (gastric mucosa, salivary glands) from HCV-negative patients and to rat liver. Positive controls of PCR amplification included performance of RT-PCR–ISH with β-actin primers and an FITC-labeled actin probe. The presence of HCV was also verified by a standard RT-PCR assay performed on homogenates obtained from frozen liver tissues of several patients.

Immunohistochemistry

Two mouse monoclonal antibodies (MAbs) corresponding to clone Tordji 32 and to a mixture of clones Tordji 22 and 32 were used (Clonatec; Paris, France). They recognize the superoxide dismutase–HCV recombinant polypeptide (Ortho

Diagnostic C100-3 Elisa kit). Sections were treated with primary MAb at 1:50 dilution for 1 hr at RT, then with secondary biotinylated horse anti-mouse IgG (Vector Labs, Burlingame, CA) and with the avidin–biotin–peroxidase complex (Vectastain ABC kit; Vector Labs). Specificity of immunoreaction was tested by omitting primary antibodies. Citrate antigen retrieval by microwave heating before IHC was performed in all cases.

Quantification of In Situ RT-PCR Results and Statistical Analysis

For each biopsy, counting of HCV labeled vs unlabeled hepatocytes was done in 10 microscopic fields at × 400 magnification. Counts were expressed as percentage of labeled hepatocytes. The surface of the specimens was roughly estimated by measuring width and length. The number of lymphoid follicles in each biopsy was counted and was expressed per 10 mm^2. Counts were made by two independent observers. All values were expressed as mean ± 1 SEM. Differences in results between responders vs non-responders and different times (B1, B2, B3) were statistically analyzed by the nonparametric Mann–Whitney *U*-test for independent populations. Results relative to the same subjects were compared by the paired data test. Level of significance was set at $p<.05$.

Results

Histology

Histological features proved the chronic hepatitis with portal nodular lymphoid infiltration and patchy necrosis (Figure 1). Patients were classified as responders ($n = 12$) to INF-α vs non-responders ($n = 9$). The responders showed a normalization of serum ALT levels persisting 1 year after completion of INF-α; the non-responders had elevated ALT. Responders had a significantly decreased number of portal lymphoid follicles per 10 mm^2 of liver tissue in B2 and B3 biopsies compared to B1 (Table 1). Four of the five patients with all three biopsies were responders. Time-course evolution in their livers of the numbers of lymphoid follicles showed the same significant variations as those observed in all responders (Figure 2).

Serum RT-PCR for HCV

RT-PCR results were clear-cut (Figure 3). At B1, all patients were seropositive. At B2, no responders had measurable HCV copies per ml by bDNA assay, but one patient was seropositive by nested RT-PCR. In the non-responders, one was transiently seronegative. At B3, all responders were seronegative and all non-responders were seropositive with relative increased copy numbers compared to B2.

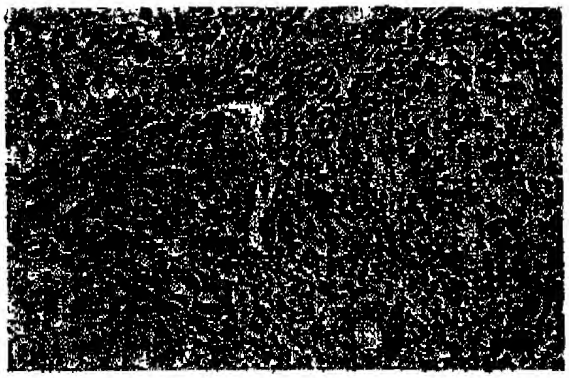

Figure 1 Histological alterations in liver biopsy specimens of HCV chronic hepatitis at B1 time. The portal area contains a nodular lymphocytic infiltrate with patchy necrosis (Mayer's hemalum–eosin–saffron). Bar = 30 μm.

In Situ HCV Detection

ISH. ISH was negative with both Genset and Argene probes. A slight but inconstant signal was noted when microwave pretreatment was performed.

In Situ RT-PCR. The signal obtained by in situ RT-PCRd and by RT-PCR–ISH was identical in the 38 biopsies, regardless of primers and probes utilized (Figures 4A–4E). Localization of HCV RNA was intranuclear, with either a diffuse or a peripheral pattern. Signal was generally strong in many hepatocytes and in nuclei of bile duct cells and portal lymphocytes, although in the latter case the number of labeled nuclei was lower. A signal located on the cytoplasm was also sometimes seen, especially with RT-PCR–ISH (Figures 4A, 4B, and 4E). No difference was seen between results obtained with the two different thermocyclers.

Table 1 Evolution of HCV in situ RT-PCR signal in hepatocytes and lymphoid follicles in responders and non-responders in the course of IFN-α treatment[a]

	No. RT-PCR-positive hepatocytes (%)			No. lymphoid follicles per 10 mm^2 liver tissue		
	B1	B2	B3	B1	B2	B3
Responders	58.00 ± 9.70 ($n = 5$)	52.20 ± 5.72 ($n = 9$)	61.00 ± 5.86 ($n = 10$)	4.40 ± 1.26 ($n = 5$)	0.88[b] ± 0.35 ($n = 9$)	0.68[c] ± 0.29 ($n = 10$)
Non-responders	60.00 ± 10.00 ($n = 3$)	55.00 ± 14.40 ($n = 4$)	75.70 ± 7.83 ($n = 7$)	3.33 ± 1.20 ($n = 3$)	2.00 ± 1.00 ($n = 4$)	2.21 ± 0.70 ($n = 7$)

[a]Statistics: vs B1 in the responders' group: [b]$p<0.004$, [c]$p<0.003$.

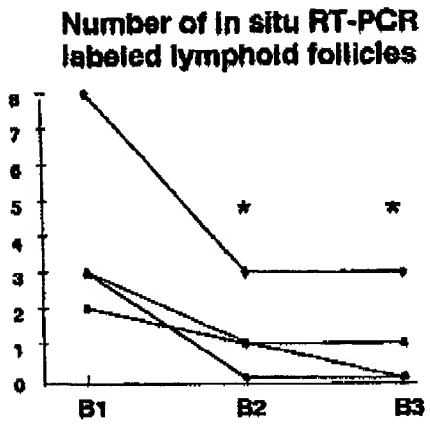

Figure 2 Numbers of lymphoid follicles in liver biopsies at B1 (before IFN-α treatment), B2 (at treatment cessation), and B3 (1 year later) in the same four patients responding to IFN-α therapy
* $p<0.05$ vs B1.

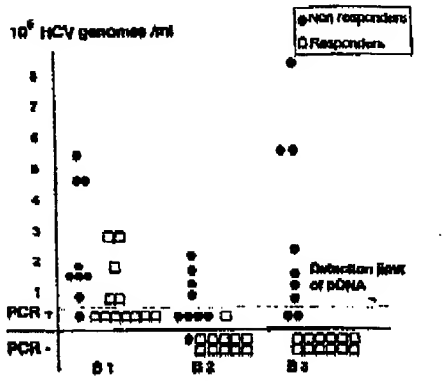

Figure 3 Serum HCV RNA levels in 21 patients according to response to interferon therapy. The bDNA quantification was possible for at least 0.35 × 10^6 viral copies/ml. Below this value, nested RT-PCR was performed, indicating only the presence or absence of viral RNA in serum.

Controls Controls for each step confirmed the specificity of signal. No signal was seen for RT-PCR-ISH and in situ RT-PCRd when RNase or DNase incubation was performed before or after reverse transcription, respectively. For RT-PCR-ISH, postamplification signal was abolished by incubation with DNase before ISH (Figure 4F). No signal was seen with heterologous probes or primers by in situ RT-PCRd or RT-PCR-ISH on HCV-seropositive patients' livers and also with homologous primers and probes on HCV-seronegative patients' livers (Figure 4G). HBV-positive liver biopsies sometimes had slight staining of the nuclear membrane with HCV probes and primers, quite different in intensity from that observed in HCV-positive liver.

Quantification of RT-PCR Signal

On the 38 hepatic biopsies studied, there was diffuse panlobular distribution of positive hepatocytes with variable staining intensity. Importantly, responders continued to have HCV nuclear signal at B2 and B3 times (Figure 5). The mean percentage of labeled hepatocytes in those patients remained constant (50–60%) and did not differ from non-responder values (Table 1). This was also checked in the four responders who underwent B1, B2, and B3 biopsies. Nevertheless, in all responders, not only did the number of lymphoid follicles decrease (Table 1) but the nuclear intensity labeling and the number of labeled nuclei in these lymphoid follicles were diminished, although no quantification was performed for this last variable. HCV RT-PCR assays were also performed on RNA extracted from frozen liver biopsies of four patients at B3. All these patients, including one responder to IFN-α therapy, appeared positive for HCV.

Immunohistochemical Results

The results were similar regardless of the antibody used. Treatment of slides by microwave heating did not improve signal specificity. The localization of signal, as with in situ amplification RT-PCR, was intranuclear in hepatocytes, bile duct cells, and portal lymphocytes (Figures 6A–6C) and occasionally cytoplasmic (Figure 6C). However, the immunoreactivity in hepatocyte nuclei was sometimes very weak (Figure 6B), and the number of labeled hepatocytes was always less than that seen with in situ RT-PCR. In addition, IHC was relatively insensitive because half of the samples from HCV livers at B1 were negative (Figure 5), despite the presence of HCV signal by in situ RT-PCR. It is noteworthy that IHC positivity also persisted at B2 and B3 in most responders. No signal was observed in heterologous tissue or in liver tissue from HCV-negative patients used as controls.

Discussion

Our results are in agreement with previous findings describing viral localization at the perinuclear membrane (Lidonnici et al. 1995) or within the nuclei of hepatocytes (Haruna et al. 1993; Nouri-Aria et al. 1993). We noted with RT-PCR-ISH, but rarely with in situ RT-PCRd, as already observed by Nouri-Aria et al. (1993) and occasionally by Nuovo (1994), a diffuse and inconstant staining within the cytoplasm, always associated with the nuclear signal. However, in our opinion, the specificity of the cytoplasmic signal is

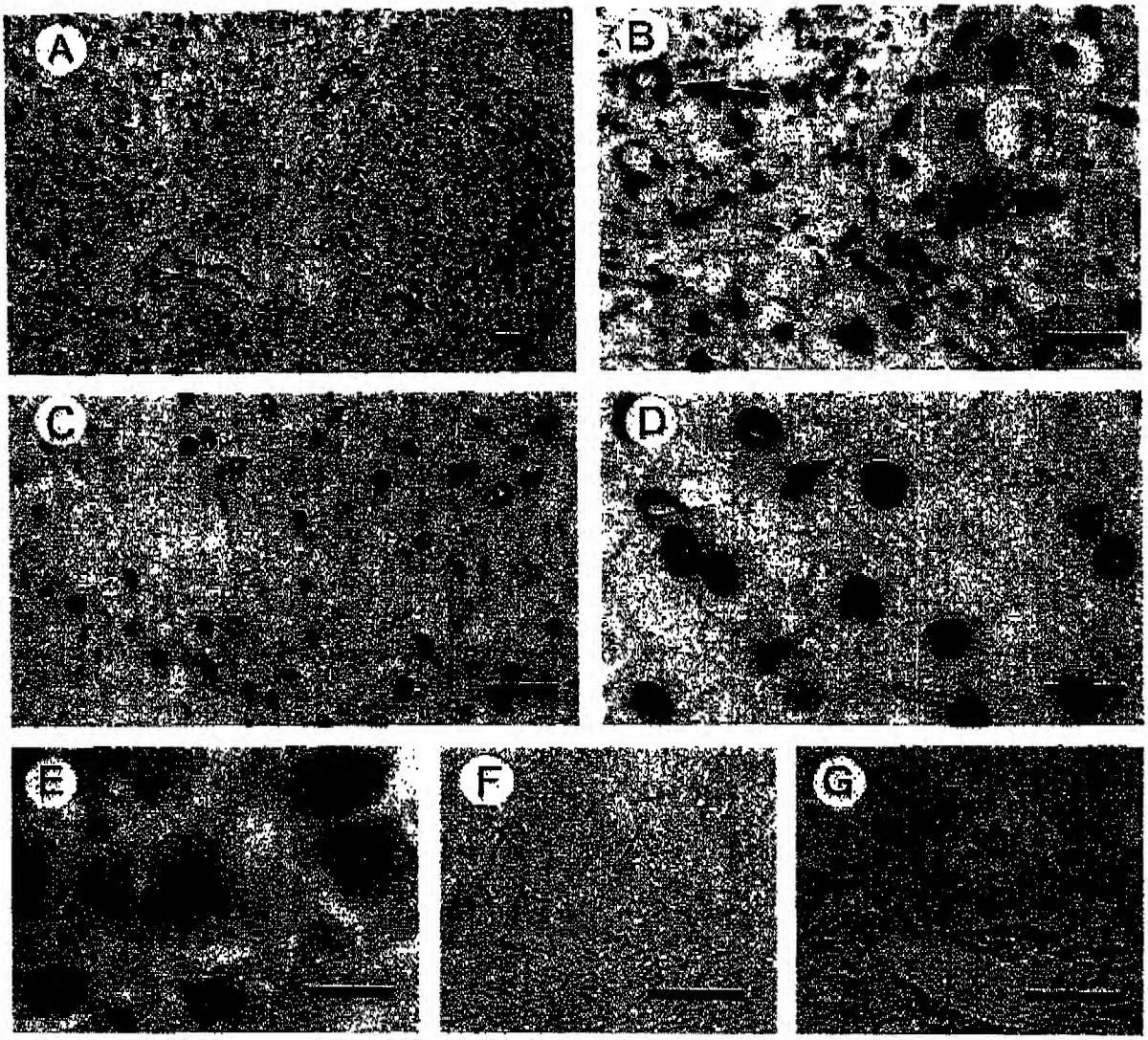

Figure 4 (A) Liver biopsy of a non-responder patient at the end of IFN-α treatment (B2) after RT-PCR–ISH with FITC-labeled probe, at low magnification. In this strongly positive area, intranuclear signal is seen in many hepatocytes, lymphocytes, and bile duct cells (no counterstaining). Bar = 30 μm. (B) Same liver biopsy at higher magnification. The signal is evident over the nuclei of hepatocytes, bile duct cells (arrow), and lymphoid cells, and also within the cytoplasm of hepatocytes (no counterstaining). Bar = 30 μm. (C) In situ RT-PCRd with FITC-labeled primers 1 year after treatment cessation (B3) in a responder's liver. Strong nuclear signal is present in a diffuse or marginal pattern. No cytoplasmic signal is seen (no counterstaining). Bar = 30 μm. (D) Same liver biopsy at higher magnification. Bar = 10 μm. (E) In situ RT-PCRd in another responder's liver at B3. Signal is consistently intranuclear but weaker cytoplasmic staining is also evident (no counterstaining). Bar = 10 μm. (F) Liver biopsy from a non-responder at B1. In situ RT-PCR–ISH signal is abolished with postamplification DNase incubation (no counterstaining). Bar = 30 μm (G) Liver biopsy from an HCV-negative patient with alcoholic cirrhosis. After in situ RT-PCRd, no signal is visible (Mayer's hemalum–eosin). Bar = 30 μm.

rather doubtful and could be due to some background. Nevertheless, it was not possible to totally eliminate it in spite of many high-stringency washings. Some authors (Komminoth et al. 1994) stated that the in situ RT-PCRd is not reliable and the nuclear signal not specific because in their control experiments, omitting primers yielded positive nuclear signals similar to those they observed with primers. In our experiment, omitting primers indeed yielded negative results and all controls confirmed the specificity of the signal

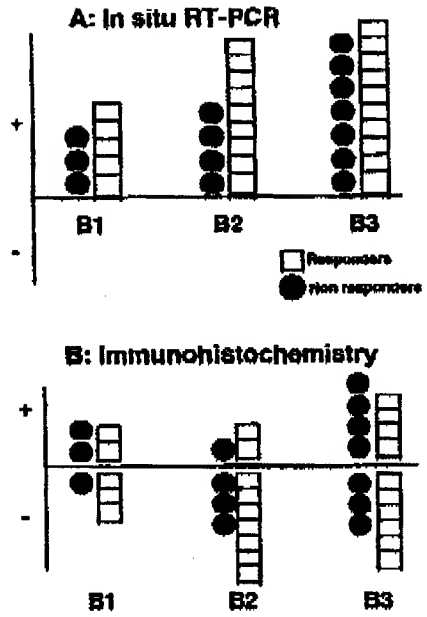

Figure 5 Diagram showing the number of positive patients for HCV detection by in situ RT-PCR (A) and IHC (B) at B1, B2 and B3. By in situ RT-PCR, HCV detection was positive at all time points for responders and non-responders. +, positive signal; -, negative signal.

reaction. It must be pointed out that Nuovo (1994) obtained the same perinuclear signal with a direct incorporation of digoxigenin-labeled nucleotides during the in situ amplification and with primers other than ours.

In the literature, HCV was found mainly in hepatocytes (Negro et al. 1992; Haruna et al. 1993; Nouri-Aria et al. 1993, Lau et al. 1994; Gastaldi et al. 1995) but also in Kupffer cells (Nuovo 1994), mononuclear cells (Nouri-Aria et al. 1993; Lau et al. 1994), sinusoidal cells (Nouri-Aria et al. 1993) and bile duct cells (Nouri-Aria et al. 1993), although rarely (Nuovo 1994). Our own results confirm most of these localizations. Unlike Park et al. (1996), who obtained inconstant results by performing RT-PCR assay after HCV RNA extraction from formol-fixed, paraffin-embedded liver biopsies, we had no problem of sensitivity with the in situ RT-PCR directly realized in such liver tissues. These authors postulated that formalin-induced cross-links may inhibit extension by reverse transcriptase. We also suggest that RNA extraction after dewaxing at high temperature could induce a large loss of HCV RNA copies, leading to false-negative results.

As already reported by Chamlian et al. (1993), the first results obtained by IHC with antisera raised against either HCV structural and nonstructural pro-

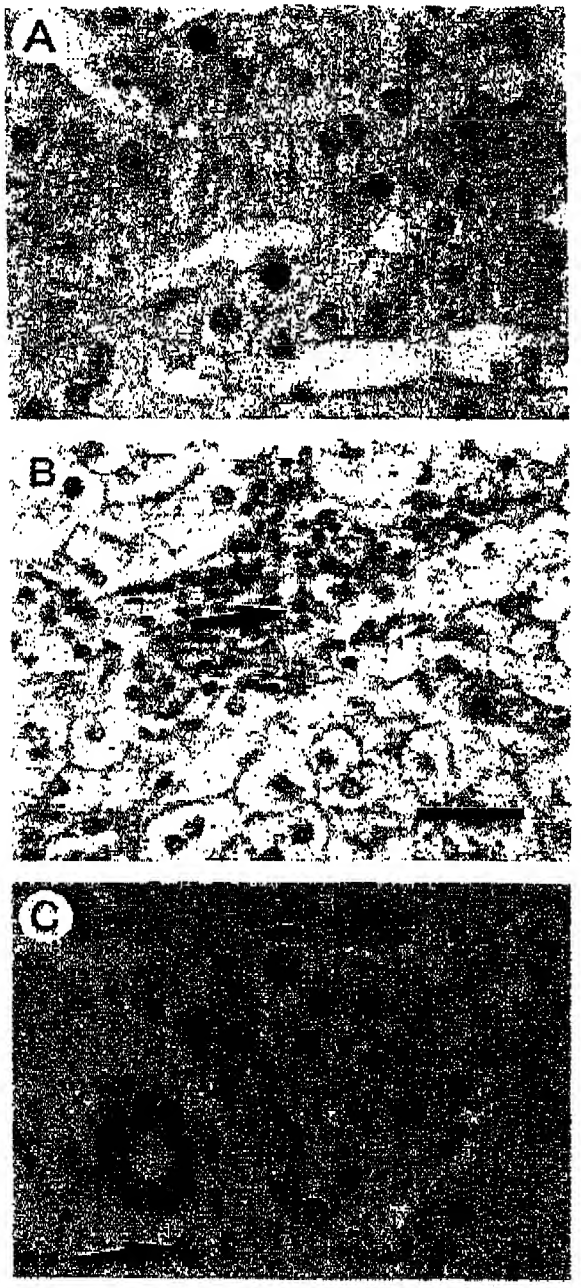

Figure 6 IHC with monoclonal antibodies against the superoxide dismutase-HCV recombinant polypeptide (C100-3 protein) clones Torq 22 and 32 (counterstaining with Mayer's hematium). (A) A nuclear signal is seen in the nuclei of two hepatocytes in a responder. (B) Nuclear signal in few hepatocytes and in bile duct cells (arrow). (C) Liver biopsy of a non-responder patient at B1. The nuclear signal is visible in relatively numerous nuclei of hepatocytes, bile ductule cells, and in lymphocytes (arrow). Diffuse cytoplasmic staining is also evident in some hepatocytes and in all duct cells. Bar = 30 μm.

teins, core antigen, E1, E2, envelope or NS2, NS3, NS4 and NS5 antigens showed a cytoplasmic signal (Hiramatsu et al. 1992, Tsutsumi et al. 1994; Uchida et al. 1994, Nouri-Aria et al. 1995; Barba et al. 1997). When using the two MAb clones Tordjt 22 and 32, which are currently the HCV antisera commercially available (Clonatec), Chamlian et al. (1993) obtained exclusively a nuclear signal in hepatocytes. We confirmed this signal but also noted some diffuse cytoplasmic staining in liver samples displaying higher numbers of immunoreactive hepatocytes. All control specimens were negative. However, IHC was less reliable than our in situ RT-PCR techniques.

In this work, HCV RNA was studied for the first time by in situ RT-PCR in infected patients' livers before and after IFN-α therapy. If our findings reflect the true HCV localization, they indicate the persistence of HCV latency within responder livers, even 1 year after IFN-α therapy cessation. Although the number of in situ RT-PCR-positive hepatocytes stayed quite constant in responders (about 50–60%), the number of lymphoid follicles decreased, as well as the HCV nuclear positivity in the lymphocytes. The two latter facts suggest that, as in other forms of chronic hepatitis, the lymphoid population in the liver is one of the factors that mediates disease activity (Navas et al. 1994) or relapse of the disease (Moldvay et al. 1994). Several hypotheses must be raised to explain the discrepancy between serum RT-PCR and in situ RT-PCR results: (a) a lack of sensitivity of serum RT-PCR (level detection 10^2 HCV genomes/ml) while in situ RT-PCR could evidence a very low viral copy number per cell, such as one or two HPV copies in Siha cells (Nuovo 1994), or (b) a level of viral replication in liver too low to be detected in serum. This viral form, probably in a latent state, may explain the relatively high chronic hepatitis C relapse. Some authors (Saleh et al. 1994) report negative RT-PCR assay in frozen liver tissues from responders to IFN-α therapy. We believe that the conditions of the cryoprotection, the very low level of viral copy numbers, and the presence in tissue of enzymes inhibiting the reaction may lead to false-negative results. In our experiment, RT-PCR assay performed on quickly extracted HCV RNA after liver biopsy was positive in one responder at B3. Certainly, our findings need to be confirmed by further investigations. In addition, we cannot rule out the possibility that the viral persistence is transient and that the virus does disappear in responder livers after a longer time following the IFN-α therapy cessation.

Acknowledgments

Supported by the ARC, grant number 2003.

We thank Pr C. Degott and Pr P. Marcellin for having given us the opportunity to study HCV-infected liver biopsies and M. Martinot-Peignoux for performing HCV RNA RT-PCR in serum.

Literature Cited

Barba G, Harper F, Harada T, Kohara M, Goulinet S, Matsuura Y, Eder G, Schaff Z, Chapman MJ, Miyamura T, Brechot C (1997) Hepatitis C virus core protein shows a cytoplasmic localization and associates to cellular lipid storage droplets. Proc Natl Acad Sci USA 94:1200–1205

Chamlian A, Benkoel L, Ikoli JF, Brisse J, Jacob T (1993) Nuclear immunostaining of hepatitis C infected hepatocytes with monoclonal antibodies to C100-3 nonstructural protein. Comparison of immunogold silver staining with other immunohistochemical methods. Cell Mol Biol 39:243–251

Feray C, Samuel D, Thiers V, Gigou M, Pichon F, Bismuth A, Reynes M, Maisonneuve P, Bismuth H, Brechot C (1992) Reinfection of liver graft by hepatitis C virus after liver transplantation. J Clin Invest 89:1361–1365

Castaldi M, Massacrier A, Planells R, Robaglia-Schlupp A, Portal-Bartolomei I, Bourliere M, Quilici F, Fiteni J, Mazzella E, Cau P (1995) Detection by in situ hybridization of hepatitis C virus positive and negative RNA strands using digoxigenin-labeled cRNA probes in human liver cells. J Hepatol 23:509–518

Haruna Y, Hayashi N, Hiramatsu N, Takehara T, Hagiwara H, Sasaki Y, Kasahara A, Fusamoto H, Kamada T (1993) Detection of hepatitis C virus RNA in liver tissues by an in situ hybridization technique. J Hepatol 18:96–100

Hiramatsu N, Hayashi N, Haruna Y, Kasahara A, Fusamoto H, Mori C, Fuke I, Okayama H, Kamada T (1992) Immunohistochemical detection of hepatitis C virus-infected hepatocytes in chronic liver disease with monoclonal antibodies to core, envelope and NS3 regions of the hepatitis C virus genome. Hepatology 16:306–311

Houghton M (1996) Hepatitis C Viruses. In Fields BN, Knipe DM, Howley PM, Chanock RM, Melnick JL, Monath TP, Roizman BR, Straus SE, eds. Field's Virology 3rd ed. Philadelphia, Lippincott-Raven Publishers, 1035–1058

Komminoth P, Adams V, Long A, Roth J, Saremaslani P, Flury R, Schmid M, Heitz PhU (1994) Evaluation of methods for Hepatitis C virus detection in archival liver biopsies. Pathol Res Pract 190:1017–1025

Lau GKK, Fang JWS, Wu PC, Davis GL, Lau JYN (1994) Detection of hepatitis C virus genome in formalin-fixed paraffin-embedded liver tissue by in situ reverse transcription polymerase chain reaction. J Med Virol 44:406–409

Lidonnici K, Lane B, Nuovo GJ (1995) Comparison of serologic analyses and in situ localization of PCR-amplified cDNA for the diagnosis of hepatitis C infection. Diagn Mol Pathol 4:98–107

Moldvay J, Deny P, Pol S, Brechot C, Lamas E (1994) Detection of hepatitis C virus RNA in peripheral blood mononuclear cells of infected patients by in situ hybridization. Blood 83:269–273

Navas S, Castillo I, Bartolome J, Marriott E, Herrero M, Carreno V (1994) Positive and negative hepatitis C virus RNA strands in serum liver and peripheral mononuclear cells in anti-HCV patients: relation with liver lesions. J Hepatol 21:182–186

Negro F, Pacchioni D, Shimizu Y, Miller Rh, Bussolati G, Purcell Rh, Bonino F (1992) Detection of intrahepatic replication of hepatitis C virus RNA by in situ hybridization and comparison with histopathology. Proc Natl Acad Sci USA 89:2247–2251

Nouri-Aria KT, Sallie R, Mizokami M, Portmann BC, Williams R (1995) Intrahepatic expression of hepatitis C virus antigens in chronic liver disease. J Pathol 75:77–83

Nouri-Aria KT, Sallie R, Sangar D, Graeme JMA, Smith H, Byrne J, Portmann B, Eddleston ALWF, Williams R (1993) Detection of genomic and intermediate replicative strands of hepatitis C virus in liver tissue by in situ hybridization. J Clin Invest 91:2226–2234

Nuovo GJ (1994) PCR In Situ Hybridization. Protocols and Applications. 2nd ed. New York, Raven Press

Park YN, Abe K, Li H, Hsuih T, Thung SN, Zhang DY (1996) De-

tection of hepatitis C virus RNA using ligation-dependent polymerase chain reaction in formalin-fixed, paraffin-embedded liver tissues. Am J Pathol 149:1485–1491

Saleh MG, Tibbs CJ, Koskinas J, Pereira LMMB, Bomford AB, Portmann BC, McFarlane IG, Williams R (1994) Hepatic and extrahepatic hepatitis C virus replication in relation to response to interferon therapy. Hepatology 20:1399–1404

Tsutsumi M, Urashima S, Takada A, Date T, Tanaka Y (1994) Detection of antigens related to hepatitis C virus RNA encoding the NS5 region in the livers of patients with chronic type C hepatitis. Hepatology 19:265–272

Uchida T, Shikata T, Tanaka E, Kiyosawa K (1994) Immunoperoxidase staining of hepatitis C virus in formalin-fixed paraffin-embedded needle liver biopsies. Virchows Arch 424:465–469

Walker F, Bedel C, Dauge-Geffroy MC, Lehy T, Madelenat P, Potet F (1996) Improved detection of human papillomavirus infection in genital intraepithelial neoplasia in human immunodeficiency virus positive (HIV+) women by polymerase chain reaction-in situ hybridization. Diagn Mol Pathol 5:136–146

Weiner AJ, Kuo G, Bradley DW, Bonino F, Saracco G, Lee C, Rosenblatt J, Choo QL, Houghton M (1990) Detection of hepatitis C viral sequences in non-A, non-B hepatitis. Lancet 335:1–3

nppena: x D
USSN 09/719,277

THE JOURNAL OF BIOLOGICAL CHEMISTRY
© 2003 by The American Society for Biochemistry and Molecular Biology, Inc
Vol. 278, No. 42, Issue of October 17, pp. 40503–40513, 2003
Printed in U.S.A.

Two Alternative Translation Mechanisms Are Responsible for the Expression of the HCV ARFP/F/Core+1 Coding Open Reading Frame*

Received for publication, May 27, 2003, and in revised form, July 21, 2003
Published, JBC Papers in Press, July 21, 2003, DOI 10.1074/jbc.M305504200

Niki Vassilaki and Penelope Mavromara‡

From the Molecular Virology Laboratory, Hellenic Pasteur Institute, 127 Vas. Sofias Avenue, Athens, Greece 11521

HCV-1 produces a novel protein, known as ARFP, F, or core+1. This protein is encoded by an open reading frame (ORF) that overlaps the core gene in the +1 frame (core+1 ORF). *In vitro* this protein is produced by a ribosomal frameshift mechanism. However, similar studies failed to detect the ARFP/F/core+1 protein in the HCV-1a (H) isolate. To clarify this issue and to elucidate the functions of this protein, we examined the expression of the core+1 ORF by the HCV-1 and HCV-1a (H) isolates *in vivo*, in transfected cells. For this purpose, we carried out luciferase (LUC) tagging experiments combined with site-directed mutagenesis studies. Our results showed that the core+1-LUC chimeric protein was efficiently produced *in vivo* by both isolates. More importantly, neither changes in the specific 10-A residue region of HCV-1 (codons 8–11), the proposed frameshift site for the production of the ARFP/F/core+1 protein *in vitro*, nor the alteration of the ATG start site of the HCV polyprotein to a stop codon significantly affected the *in vivo* expression of the core+1 ORF. Furthermore, we showed that efficient translation initiation of the core+1 ORF is mediated by internal initiation codon(s) within the core/core+1-coding sequence, located between nucleotides 583 and 606. Collectively, our data suggest the existence of an alternative translation initiation mechanism that may result in the synthesis of a shorter form of the core+1 protein in transfected cells.

Hepatitis C virus (HCV)[1] is the main etiologic agent of post-transfusion and sporadic non-A, non-B hepatitis in the world (1). This virus establishes chronic infection in most acutely infected individuals, frequently leading to liver cirrhosis and hepatocellular carcinoma (2, 3). HCV is an enveloped, single-stranded, positive sense RNA virus. It is a member of the Hepacivirus genus within the Flaviviridae family (4). The viral genome is ~10-kb long and encodes a 3011-amino acid polyprotein precursor. This polyprotein is co- and post-translationally processed by cellular and viral proteases giving rise to three structural (core, E1, and E2) and at least six non-structural (NS2, NS3, NS4A, NS4B, NS5A, and NS5B) proteins (5). Initiation of translation of the HCV genome is controlled by an internal ribosome entry site (IRES) located mainly within the 5′-non-coding region of the viral RNA, between nucleotides 42 and 341 or 356, the 3′-limit being controversial (6). The core protein, which forms the viral nucleocapsid, is predicted to be 191 amino acids long and to have a molecular mass of 23 kDa (p23). Further processing of p23 produces the mature core protein (p21), consisting of 173–182 amino acids (7, 8).

An additional HCV polypeptide of 16/17 kDa (p16/p17) has recently been discovered (9, 10, 11). This protein is encoded by the open reading frame (ORF) that overlaps the core gene in the +1 frame (core+1 ORF). This polypeptide, named the ARFP (for alternative reading frame protein), F (for frameshift protein) or core+1 (to describe the location of this novel protein), is synthesized *in vitro* from the initiator codon of the polyprotein sequence followed by a +1 ribosomal frameshift operating in the region of core codons 8–14 (9, 11). Moreover, circulating anti-core+1 antibodies have been detected in HCV-infected individuals, suggesting that this protein is produced during natural HCV infection (9, 10, 11).

However, the function of the ARFP/F/core+1 protein in the life cycle of HCV remains unknown. Interestingly, a non-structural protein is encoded by the N-terminal structural region of a number of positive sense RNA viruses like Pestiviruses (N^{pro} protease, Ref. 12) and Apthoviruses (L* protein, Ref. 13). These proteins are believed to play essential roles in the viral life cycle. Moreover, it is tempting to speculate that some of the functional properties attributed to the core are in fact due to the ARFP/F/core+1 protein, as both proteins are probably produced simultaneously in certain experimental conditions (7, 8). To investigate the biological importance of the ARFP/F/core+1 protein, as a first step, we assessed the expression of the core+1-coding sequence in mammalian cells. During the course of these experiments, we examined the expression of the core+1 ORF by HCV-1, which efficiently synthesizes the ARFP/F/core+1 p16/p17 protein *in vitro*, in rabbit reticulocyte lysates, and by HCV-1a (H), in which no p16/p17 has been detected (9). For this, we carried out transient transfection assays based on the luciferase tagging approach combined with site-directed mutagenesis analysis. The luciferase tagging approach was chosen because it allows the sensitive enzymatic detection of the core+1 translation product and can be used to assess the relative expression levels of the core and core+1 ORFs in the two HCV isolates.

Our results show that, unlike in the *in vitro* expression studies, both HCV-1 and HCV-1a (H) efficiently express the core+1 ORF in transfected cells. More importantly, our generic analysis and immunoprecipitation experiments suggest that in transfected cells efficient translation initiation of core+1 is mediated from internal translation initiation codons located between nt 583 and 606, resulting in the synthesis of a shorter form of the core+1-LUC chimeric protein. This form of the

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed: Molecular Virology Laboratory, Hellenic Pasteur Institute, 127 Vas. Sofias Ave., Athens, Greece 11521. Tel.: 210-6478875; Fax: 210-6478877; E-mail: penelopm@hol.gr.

[1] The abbreviations used are: HCV, hepatitis C virus; ORF, open reading frame; nt, nucleotides; IRES, internal ribosome entry site; LUC, luciferase; CAT, chloramphenicol acetyltransferase

TABLE I
List of oligonucleotides and constructs used in the mutational analysis

	Mutation	Oligonucleotides (sense)	*In vitro* plasmids	*In vivo* plasmids
WT: pHPI-1333/HCV-1/CAT-IRES-core-LUC				
	Mutations			
	N1	5′-$^{-45a}$ CAGGGGCCCTAGATTAGGTGTGCGCGCGAC-3′	pHPI-1342	pHPI-1342
	N3	5′-$^{-40a}$ CGTGCACCTAGAGCACGGATCCTAAACCTC-3′	pHPI-1343	pHPI-1343
	N6	5′-$^{-30a}$ CGTCAAGTTCTAAGGTGCCGGTC-3′	pHPI-1344	pHPI-1344
	N18	5′-$^{-30a}$ CCTAAACCTCAAGGAAAAACCAAACGTAACACC-3′	pHPI-1382	pHPI-1382
	N19	5′-$^{-40a}$ CCTAAACCTCAAAGACAAACCAAACGTAACACC-3′	pHPI-1383	pHPI-1383
	N21	5′-$^{-57a}$ CTCAGCCCGTAGACCCTTGGCC-3′	pHPI-1380	pHPI-1380
	N22	5′-$^{-50a}$ CTCTATGGCATAGAGGGCTGCGGGT-3′	pHPI-1381	pHPI-1381
	N23	5′-$^{-54a}$ TTGGCCCCTCTGGGGCAATGAGGGC-3′	pHPI-1399	pHPI-1399
	N24	5′-$^{-50a}$ CTCTATGGCAGGGAGGGCTGCGGGT-3′	pHPI-1400	pHPI-1400
	N25	5′-$^{-50a}$ TTGGCCCCTCTGGGGCAGGGAGGGCTGCGGGT-3′	pHPI-1401	pHPI-1401
NC:pHPI-1332				
WT: pHPI-1335/HCV-1a (H)/ CAT-IRES-core-LUC				
	Mutations			
	N1	5′-$^{-45a}$ CAGGGGCCCTAGATTAGGTCTGCGCGCGAC-3′	pHPI-1345	pHPI-1345
	N3	5′-$^{-40a}$ CGTGCACCTAGAGCACGGATCCTAAACCTC-3′	pHPI-1346	pHPI-1346
	N6	5′-$^{-30a}$ CGTCAAGTTCTAAGGTGGCGGTC-3′	pHPI-1347	pHPI-1347
	N15	5′-$^{-40a}$ CCTAAACCTCAAGGAAAAACCAAACGTAACACC-3′	pHPI-1395	pHPI-1395
	N16	5′-$^{-30a}$ CCTAAACCTCAAAGACAAACCAAACGTAACACC-3′	pHPI-1396	pHPI-1396
	N21	5′-$^{-57a}$ CTCAGCCCGTAGACCCTTGGCC-3′	pHPI-1398	pHPI-1398
	N22	5′-$^{-50a}$ CTCTATGGCATAGAGGGTTGCGGGT-3′	pHPI-1397	pHPI-1397
NC: pHPI-1336				

ARFP/F/core+1 chimeric protein is synthesized even when the core is not produced.

EXPERIMENTAL PROCEDURES

Plasmid Construction and Site-directed Mutagenesis—Site-directed mutagenesis was carried out using the Quikchange™ kit (Stratagene). The templates and oligonucleotides used in the mutational analysis and the corresponding mutants are listed in Table I. All mutations were confirmed by sequence analysis. The HCV-1 cDNA sequences were obtained from plasmids path 10/17–98 and path 11/36–27, kindly provided by Dr. M. Beach (CDC). The HCV-1a (H) cDNA sequence was obtained from pDNA-C1, kindly provided by Dr. G. Inchauspé

The dicistronic constructs pHPI-1331, −1333, and −1332 contain the chloramphenicol acetyltransferase (*CAT*) gene as the first cistron followed by the entire IRES and part of the wild-type core-coding sequences (nt 9–630) from the prototype HCV-1 isolate fused to the firefly LUC gene in the 0, +1, and −1 frames, respectively. They were produced by site-directed mutagenesis from dicistronic pHPI-1314, −1318 and −1312, respectively, using primers 5′-TGGATCCAAGGGGAAGACGCC-3′ (sense) and 5′-GGCGTCTTCCCCTTGGATCCA-3′ (antisense). This set of primers converts the start codon of the luciferase-coding region (ATG) into a glycine codon (GGG). pHPI-1311, -1313, and -1312 were constructed by replacing the 203-bp *Nhe*I-*Xba*I fragment of the dicistronic pHPI-1046 (previously described, Ref. 14), carrying nt 249–407 of the IRES-core sequences fused with part of the LUC gene (the first 50 nt), with the 435-bp *Nhe*I-*Xba*I fragments of pHPI-766, -767, -768 (previously described, Ref 9), containing nt 249–630 of the IRES-core sequences fused with the first 50 nt of the LUC gene. The dicistronic constructs pHPI-1334, -1335, and -1336 carry the entire IRES and part of the wild-type core-coding region (nt 9–630) from HCV-1a (H) fused to the LUC gene in all three frames (0, +1, and −1 respectively). They were derived by site-directed mutagenesis using pHPI-1328, -1329, and -1330, respectively as templates and primers 5′-TGGATCCAAGGGGAAGACGCC-3′ (sense) and 5′-GGCGTCTTCCCCTTGGATCCA-3′ (antisense). The primers change the initiator codon of the luciferase-coding region into a glycine codon (ATG → GGG). pHPI-1328, -1329, and -1330 were generated by replacing the 203-bp *Nhe*I-*Xba*I fragment of pHPI-1046 with the 435-bp *Nhe*I-*Xba*I fragments of pHPI-748, -749, -750, respectively (previously described, Ref. 9). To facilitate plasmid characterization, most of the mutations inserted into the dicistronic constructs pHPI-1333 (HCV-1) and pHPI-1335 (HCV-1a (H)) were recloned into pHPI-1046 by replacing the wild-type 203-bp *Nhe*I-*Xba*I fragment with the corresponding fragment of the mutated template.

The monocistronic constructs pHPI-1362 and −1363 contain the same IRES-core-LUC cassette as pHPI-1333 and −1332 respectively, cloned between the *Hind*III and *Sal*I sites of pEGFPN3 (Clontech).

In Vitro Transcription and Translation—For all the plasmids, Flexi rabbit reticulocyte lysates (Promega) supplemented with 120 mM KCl and 0.5 mM $Mg(OAc)_2$ were used. DNA (3 μg) from each plasmid was linearized and transcribed *in vitro* with T7 RNA polymerase (Promega) according to the manufacturer's instructions. Wild-type pHPI-1331, -1333, -1332, -1334, -1335, -1336, and the corresponding mutated dicistronic constructs were linearized with *Pst*I. *In vitro* translation experiments were carried out on uncapped RNAs in a total volume of 25 μl using [^{35}S]methionine (Amersham Biosciences). The translation products (5 μl) were analyzed by 12% SDS-PAGE, transferred onto nitrocellulose membranes, and visualized by autoradiography.

Cells and DNA Transfection Experiments—BHK-21 and Huh-7 cells were maintained in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (DMEM/FBS) at 37 °C in a 5% CO_2 incubator. Cells seeded in 6-well plates (60% confluence) were transfected with 1 μg of plasmid DNA in the presence of Lipofectamine plus™ reagent (Invitrogen) according to the manufacturer's protocol. The medium was replaced with new DMEM/FBS 24 h post-transfection. The cells were washed twice with phosphate-buffered saline 48 h post-transfection and lysed in 260 μl of 1× luciferase lysis buffer (Promega). Firefly LUC was quantified by mixing 20 μl of extract with 100 μl of luciferase assay reagent (Promega) and measuring the luminescence directly with a Turner TD-20/20 luminometer. In the case of the dicistronic constructs, CAT was quantified with the CAT-ELISA kit (Roche Applied Science) according to the manufacturer's instructions.

Antibodies—The goat polyclonal antibody against the firefly luciferase protein was obtained from Promega Corporation at a concentration of 1 mg/ml.

Immunoprecipitation Analysis—Thirty-six hours after transfection with pHPI-1362 or pHPI-1363, monolayers of BHK-21 cells (~10^7 cells) were metabolically labeled for 12 h with 20 μCi of [^{35}S]methionine (Amersham Biosciences) per ml of methionine-free medium supplemented with 1% fetal bovine serum. The labeled cells were rinsed with phosphate-buffered saline and lysed in 500 μl of total volume of triple detergent buffer consisting of 50 mM Tris (pH 8), 150 mM NaCl, 0.1% SDS, 1% Nonidet P-40, 0.5% sodium deoxycholate, and 100 μg/ml phenylmethylsulfonyl fluoride. Cell lysates were mixed by vortexing and centrifuged at 14,000 × *g* for 10 min at 4 °C. Clarified lysates were incubated with 10 μl of anti-LUC polyclonal antibody on a rocker overnight at 4 °C. Protein G PLUS-agarose (Santa Cruz Biotechnology) was added (20 μl) to this mixture and the reactions were incubated in the same conditions for an additional 2 h. After microcentrifugation, the agarose beads were washed three times with a buffer containing 50 mM Tris (pH 8), 150 mM NaCl, 0.1% Nonidet P-40, and 1 mM EDTA. The immunoprecipitates were subsequently resolved by 10% SDS-PAGE, transferred onto nitrocellulose membranes, and detected by autoradiography.

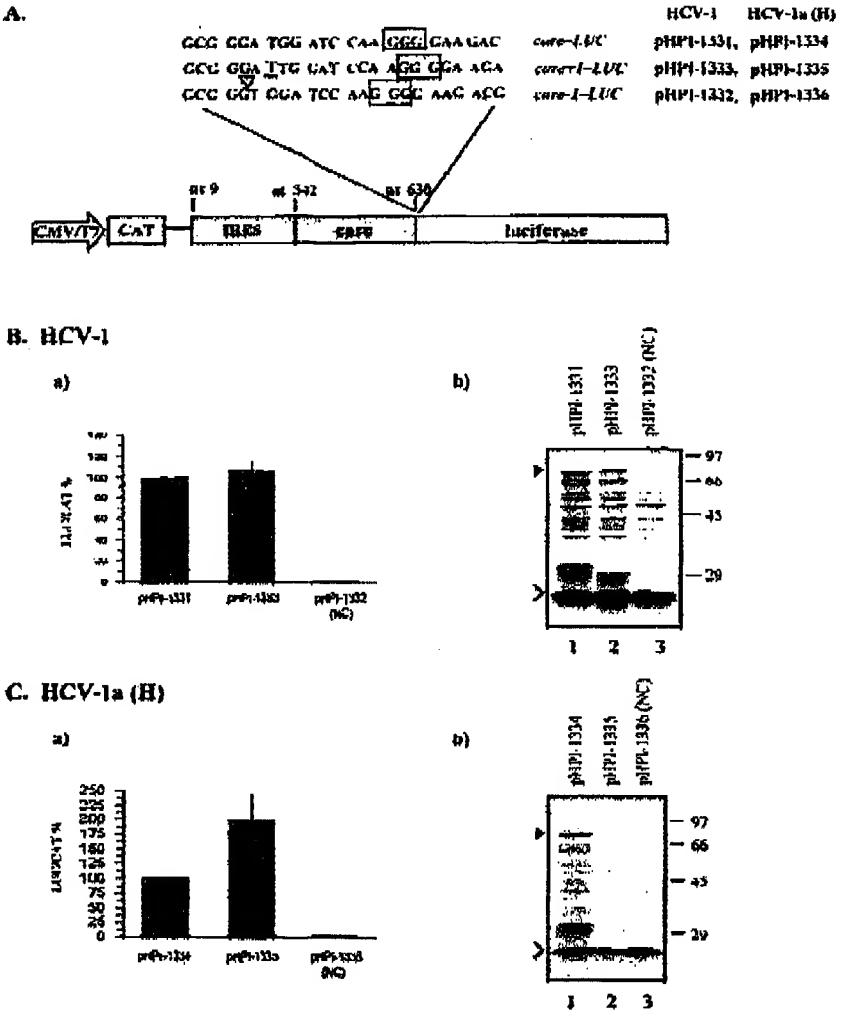

FIG. 1. **Analysis of the expression of the core+1-LUC chimeric gene.** *Panel A*, schematic representation of the CAT-LUC dicistronic constructs used for the tagging experiments. The entire IRES (nt 9–341) and part of the core-coding sequence (nt 342–630) from HCV-1 and HCV-1a (H) were fused with the LUC gene under the control of both CMV and T7 promoters of pHPI-1046 (14). The nucleotide sequences of the junction between the core and luciferase-coding regions are illustrated. The first codon of luciferase cistron, GGG, derived from the ATG initiator by site-directed mutagenesis, is *boxed*. The LUC gene was fused in the 0 frame relative to the preceding core-coding sequence in pHPI-1331 (HCV-1) and pHPI-1334 (HCV-1a (H)), in the +1 frame in pHPI-1333 (HCV-1) and pHPI-1335 (HCV-1a (H)), and in the −1 frame in pHPI-1332 (HCV-1) and pHPI-1336 (HCV-1a (H)). The *underlined* nucleotide indicates an insertion of a thymidine residue, and the inverted triangle indicates a deletion of an adenine residue. *Panels B and C, in vivo* (*a*) and *in vitro* (*b*) expression of the HCV-1 (B) and HCV-1a (H) (C) fusion constructs. *a*, duplicate cultures of BHK-21 cells were transfected with each construct and the relative ratio of LUC activity to CAT quantity was determined. *Bars* represent the means obtained in two separate experiments in duplicate. *Error bars* represent the standard deviation. *b*, each construct was transcribed *in vitro* and equal amounts of all RNAs were translated in Flexi rabbit reticulocyte lysates. Translation products were directly separated by SDS-PAGE and analyzed by autoradiography. Fusion proteins are indicated by *filled arrowheads*. *Open arrowheads* show the CAT protein. *NC*, negative control.

RESULTS

The Core+1 ORF Is Efficiently Expressed in Transfected Cells—As all previous studies concerning the expression of the core+1 ORF have been carried out primarily in *in vitro* systems based on rabbit reticulocyte lysates, we examined core+1 translation from the HCV-1 and HCV-1a (H) isolates in mammalian cells. For this, we carried out transient transfection assays, initially in BHK-21 cells, and monitored the expression of core and core+1 ORFs by tagging the ORFs with the luciferase gene (LUC) as described before (9) but with the following modifications. The cDNA sequences containing the entire IRES and part of the core-coding sequences (nt 9–630) from HCV-1 and HCV-1a (H) fused to the LUC gene in all three frames were transferred into a dicistronic vector in which CAT was the first gene. The dicistronic cassettes CAT-IRES-core-LUC were placed under the control of a CMV/T7 promoter to allow the use of the same DNA plasmid for expression *in vivo* and *in vitro*. Furthermore, to eliminate the possibility of internal translational initiation events triggered by the initiator codon of the LUC gene, the ATG was changed to a GGG codon. The DNA constructs are illustrated in Fig. 1*A*. With this modified dicistronic expression system, the expression of the LUC gene is directly related to the expression of the fused core or core+1-coding sequences, and CAT activity serves as an internal control to standardize transfection efficiencies *in vivo* or potential variations in the transcript abundance *in vitro*.

BHK-21 cells were transfected with each construct, and 48 h later LUC and CAT activities were measured. Fig. 1, *B* and *C* show the level of LUC expression relative to CAT in the transfected cells. In the case of HCV-1, substantial amounts of luciferase were expressed from the core+1-LUC cassette of pHPI-1333, as the levels of the luciferase activity were similar to that of the core-LUC fusion protein derived from pHPI-1331 (Fig. 1*Ba*). Only background levels of luciferase activity were detected following the expression of the corresponding negative control core-1-LUC construct (pHPI-1332). Surprisingly, in contrast to *in vitro* (9), very high levels of luciferase activity were observed from construct pHPI-1335, which contains the core+1 ORF from HCV-1a (H) fused to the LUC gene. The levels were about 200% of that of the HCV-1a (H) core-LUC hybrid protein yielded from pHPI-1334 (Fig. 1*Ca*). Again, the corresponding negative control plasmid (pHPI-1336) resulted in background levels of luciferase expression. Thus, the HCV-1 core+1 LUC-tagged protein is efficiently produced *in vivo*, with

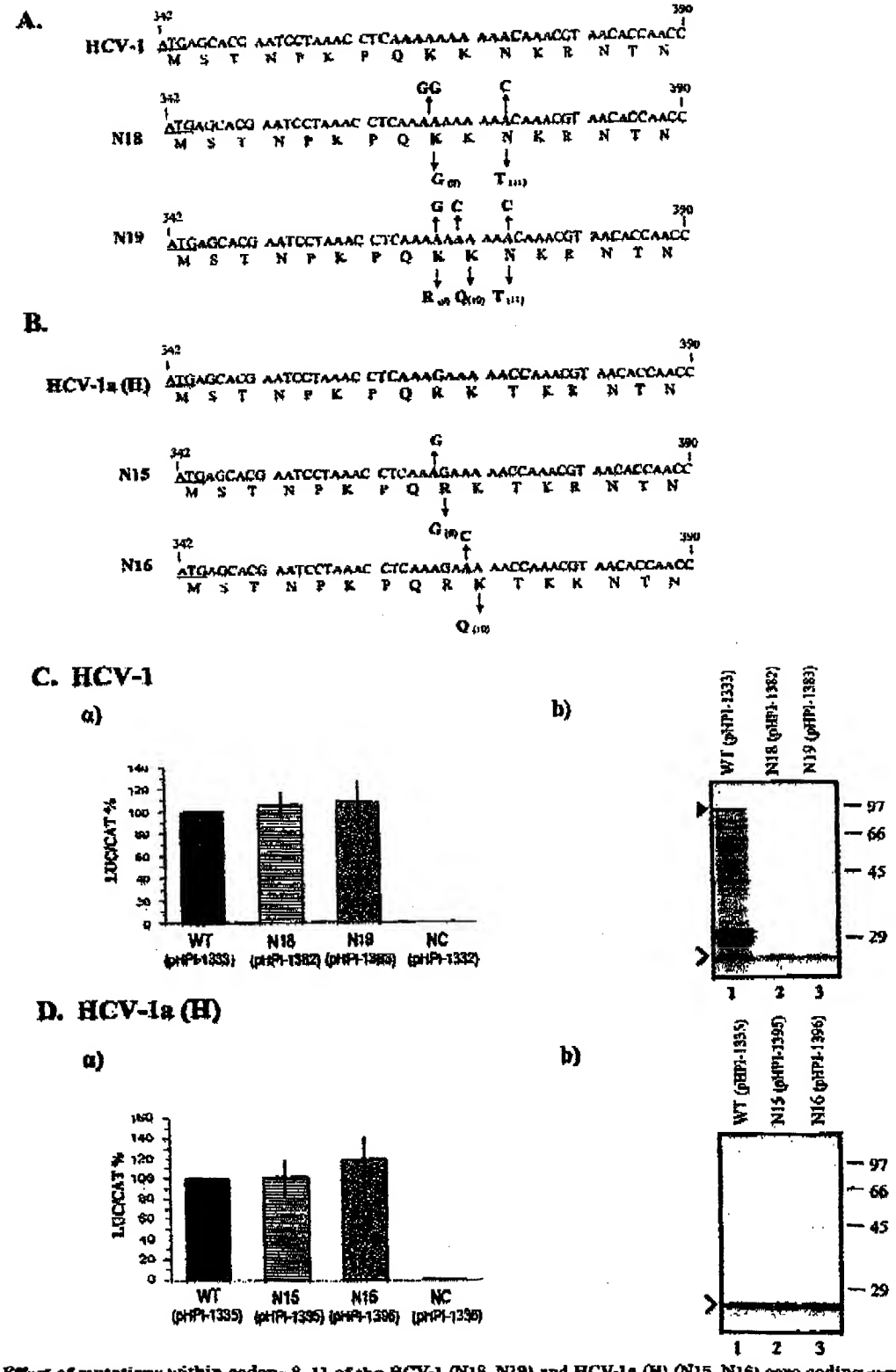

FIG. 2. **Effect of mutations within codons 8–11 of the HCV-1 (N18, N19) and HCV-1a (H) (N15, N16) core-coding sequence on the expression of core+1-LUC chimeric gene.** *Panels A* and *B*, the core nucleotide sequences in the region of codons 8–11 of the wild-type HCV-1 (*A*) and HCV-1a (H) (*B*) plasmids, as well as of the corresponding mutant variants N18, N19 (HCV-1) (*A*) and N15, N16 (HCV-1a (H)) (*B*). The

similar translation levels as the core-coding sequence. More importantly, unlike in rabbit reticulocyte lysates, our data indicate that the HCV-1a (H) isolate efficiently expresses the core+1 ORF in transiently transfected BHK-21 cells.

To confirm the divergence between the previous *in vitro* and our current *in vivo* studies, the same DNA plasmids were tested for expression *in vitro* in rabbit reticulocyte lysates (Fig. 1, *B* and *C*). As anticipated, the *in vitro* translation of the core-LUC construct from HCV-1 (pHPI-1331) or HCV-1a (H) (pHPI-1334) resulted in the synthesis of a chimeric core-LUC protein with an apparent molecular mass of 72 kDa (Fig. 1, *B* and *Cb*, *lane 1*). Furthermore, in agreement with previous data (9), when pHPI-1333, containing the core+1-LUC cassette from HCV-1, was used, a hybrid protein with an apparent molecular mass of 72 kDa was produced (Fig. 1*Bb*, *lane 2*). On the contrary, pHPI-1335, carrying the equivalent core+1-LUC cassette from HCV-1a (H), exhibited no luciferase expression (Fig. 1*Cb*, *lane 2*). The 72-kDa protein was not produced by pHPI-1332 (HCV-1) or pHPI-1336 (HCV-1a (H)), which carry the LUC gene fused to the −1 frame of the core-coding region (Fig. 1, *B* and *Cb*, *lane 3*). These data confirm previous *in vitro* expression studies from our laboratory showing that the core+1 ORF was expressed from the HCV-1 core-coding region but not from the HCV-1a (H) equivalent. They also reveal that the predominant translation mechanism that directs the expression of ARFP/F/core+1 ORF differs *in vitro* and in transfected cells.

The A-rich Sequence at Codons 8–11 of the HCV-1 Core-coding Region Is Not Essential for the Production of the ARFP/F/Core+1 Protein in Transfected Cells—The core-coding region of HCV-1a (H) lacks the 10 consecutive A residues found at codons 8–11 (nt 364–373) of the HCV-1 genome, a known slippery site prone to ribosomal frameshift events. Thus, we directly assessed the importance of the 10-A residue region for the expression of the core+1 ORF in transfected cells as compared with that in rabbit reticulocyte lysates. For this purpose, mutational studies were performed based on previously described naturally occurring mutations at the proposed frameshift site (codons 8–14). Specifically, Yeh *et al.* (15) provided evidence that a p16 protein from the core-coding region of clinical isolates from patients with hepatocellular carcinoma was produced *in vitro* in rabbit reticulocyte lysates. The detected protein was reported as being a short form of the core protein even though the protein was immunologically distinct from the p21 core. The presence of this protein was associated with specific missense mutations within codons 9–11, none of which generate a 10-A residue sequence. To determine the effect of these naturally occurring mutations on the production of the ARFP/F/core+1 protein *in vivo*, the dicistronic constructs pHPI-1333 (HCV-1) and pHPI-1335 (HCV-1a (H)) were used as templates. In the case of HCV-1 (Fig. 2*A*), insertion of mutation N18, which contains two A to G substitutions and an A to C substitution at nt 366, 367, and 373 respectively (codons 9 and 11), gave rise to pHPI-1382, whereas mutation N19, which carries an A to G substitution and two A to C substitutions at nt 367, 369, and 373 respectively (codons 9, 10, and 11), resulted in pHPI-1383. For HCV-1a (H) (Fig. 2*B*), insertion of mutation N15, which contains an A to G change at position 366 (codon 9), gave rise to pHPI-1395, and insertion of N16, which carries an A to C substitution at nt 369 (codon 10), resulted in pHPI-1396. Both N15 and N16 mutations contain a single substitution as the HCV-1a (H) isolate already carries a G and a C at positions 367 and 373 respectively. None of these mutations had a significant effect on luciferase activity *in vivo* (Fig. 2, *C* and *Da*), indicating that the presence of the 10 consecutive adenines at codons 8–11, as found in the HCV-1 isolate, is not critical for core+1 ORF expression *in vivo*. However, *in vitro* none of the above mutants produced detectable levels of the core+1-LUC hybrid protein (Fig. 2, *C* and *Db*, *lanes 2* and *3*). Thus, mutations that disrupt the 10-A residue sequence within codons 8–11 of the core-coding region fail to support the production of the core+1-LUC protein *in vitro*, whereas the same mutations have no effect on the expression of the protein *in vivo*. Thus, our data do not appear to be consistent with those previously reported (15). The reason for this is not clear at the moment, but it is likely to be due to differences in the core-coding sequence beyond codons 9–11 between the HCV isolates used in the two studies or to differences between the two rabbit reticulocyte lysate expression systems. Overall, our results indicate that the A-rich region at codons 8–11 is critical for the expression of the ARFP/F/core+1 protein only in rabbit reticulocyte lysates.

The ATG Initiator Codon of the HCV Core-coding Region Is Not Essential for the Expression of the ARFP/F/Core+1 Protein in Transfected Cells—To investigate further the molecular mechanism(s) implicated in the *in vivo* expression of the core+1 ORF, two mutations were introduced into the core-coding region of the core+1-LUC-tagged constructs of both the HCV-1 (pHPI-1333) and HCV-1a (H) (pHPI-1335) isolates (Fig. 3*A*). Mutation N3 converted the ATG initiator codon of the core ORF into a terminator codon and mutation N6 introduced a stop codon at the 25th position of the core-coding sequence at nt 414 (P_{25}, CCG). The resulting plasmids were named pHPI-1343 and pHPI-1344 for HCV-1, and pHPI-1346 and pHPI-1347 for HCV-1a (H), respectively. If a ribosomal frameshift mechanism is indeed responsible for the production of the ARFP/F/core+1 protein *in vivo*, then the N3 mutation should abrogate the production of the chimeric protein, whereas N6 should have no effect. However, transfection studies showed that not only the N3 mutation failed to block core+1 expression but also the resulting mutant exhibited increased levels of luciferase activity (Fig. 4, *A* and *B*) compared with the wild-type core+1-LUC (positive control). Similar data were obtained for both HCV-1 (Fig. 4*Aa*) and HCV-1a (H) (Fig. 4*Ba*). Additionally, the N6 mutation resulted in increased levels of luciferase activity in both isolates; the levels of luciferase activity detected in these constructs were similar to those observed in constructs carrying the N3 mutation (Fig. 4, *A* and *Ba*). These results suggest that, in contrast to *in vitro* (9), the *in vivo* translation of the core+1 ORF does not require core expression. Instead, our data suggest that blocking the translation of the core ORF has a positive effect on the translation of the core+1 ORF.

To confirm again the differences between the *in vitro* and *in vivo* results, the same DNA plasmids were tested *in vitro* in

wild-type sequences of codons 8–11 are shown in *bold*. The *arrows* indicate the inserted mutations and the *bold characters* indicate the mutated nucleotides and affected amino acids. The numbers in *brackets* indicate the number of the mutated codons. *Panels C* and *D*, the HCV-1 (*C*) and HCV-1a (H) (*D*) wild-type (pHPI-1333 and pHPI-1335, respectively) and corresponding mutants (pHPI-1382 (N18), −1383 (N19), and pHPI-1396 (N15), -1396 (N16), respectively) were used to transfect BHK-21 cells (*a*) or transcribed *in vitro* and equal amounts of RNAs were translated in Flexi rabbit reticulocyte lysates (*b*). *a*, duplicate cultures of BHK-21 cells were transfected with the wild-type or the mutated constructs. The activity of each mutant was calculated by determining the ratio of LUC activity to CAT quantity and is expressed as a percentage of that of the wild type. *Bars* represent the means observed for three separate experiments each carried out in duplicate. *Error bars* correspond to the standard deviation. *b*, 5*h* of the [^{35}S]methionine-labeled *in vitro* translation products were separated by 12% SDS-PAGE and analyzed by autoradiography. The core+1-LUC fusion protein is indicated by the *filled arrowhead*. *Open arrowheads* show the CAT protein. *WT*, wild type; *NC*, negative control.

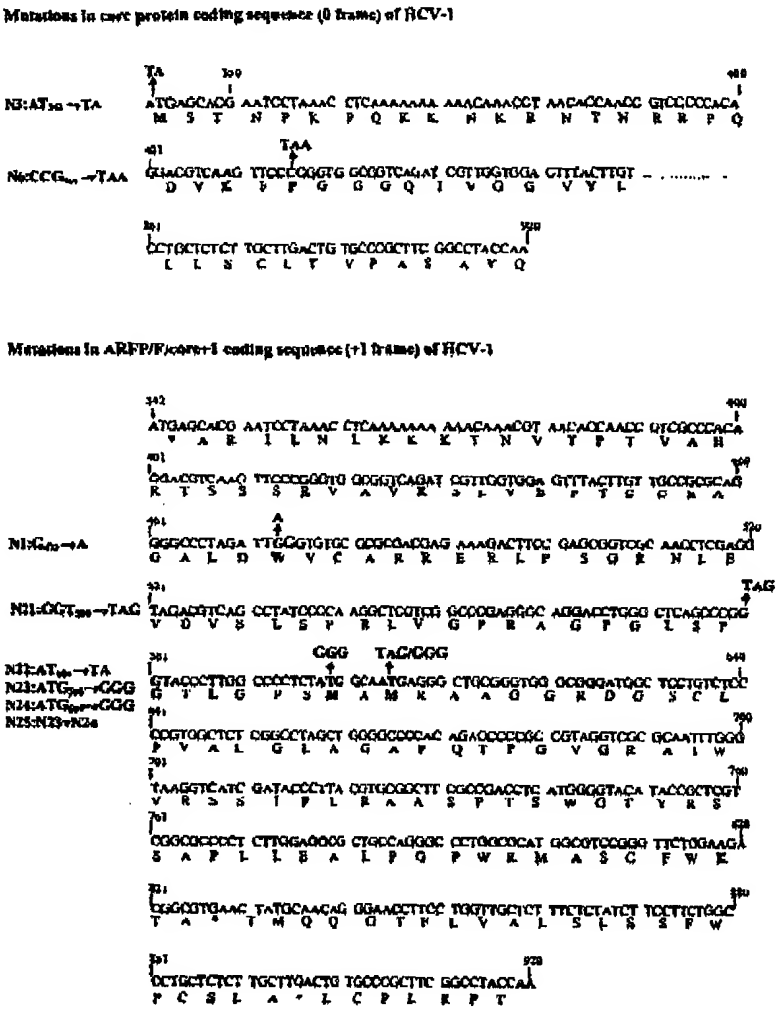

FIG. 3. **Mutational analysis of the core/core+1-coding region.** Nucleotide sequence of the HCV-1 core-coding region including mutations affecting the 0 (*A*) and +1 (*B*) open reading frames (ORFs). Inserted mutations are indicated by *arrows*. Mutated nucleotides and affected amino acids are shown in *bold*.

rabbit reticulocyte lysates. Consistent with previous *in vitro* studies (9), the N3 mutation abrogated the synthesis of the 72-kDa *core+1-LUC protein* from HCV-1 (*Fig. 4Ab, lane 2*), whereas N6 had no effect on the production of the core+1-LUC chimeric protein (Fig. 4*Ab*, *lane 3*). Furthermore, as expected, the core+1-LUC constructs (WT, N3, N6) from HCV-1a (H) failed to produce detectable levels of the chimeric protein (Fig. 4*Bb*, *lanes 1, 2, 3*). Thus, there appears to be a difference between the predominant translation mechanism for ARFP/F/*core+1* expression between rabbit reticulocyte lysates and transfected cells, as expression of the core+1 ORF *in vivo* does not require the ATG start codon of the HCV polyprotein. Taken together these results indicate that ribosomal frameshift events are not the predominant mechanism directing *core+1* expression in transfected cells.

Efficient Translation of the Core+1 ORF in Transfected Cells Is Mediated by Internal Initiation Codon(s)—As the expression of the core+1 ORF *in vivo* is not suppressed by changes in the initiator ATG or the A-rich region, we hypothesized that downstream codon(s) may function as translation initiation sites for the expression of the core+1 ORF. To examine this hypothesis two series of mutagenesis experiments were carried out.

Firstly, three nonsense mutations were separately inserted into the core+1-coding sequences of HCV-1 and HCV-1a (H), in pHPI-1333 (HCV-1) and pHPI-1335 (HCV-1a (H)), which carry the dicistronic CAT-IRES-core+1-LUC cassette. To facilitate the description of the mutations affecting the core+1 ORF, we arbitrarily defined the GCA alanine codon at nt 346 as the first codon of the core+1 ORF (Fig. 3*B*). The N1 mutation introduced a TAG stop codon into the core+1 ORF at nt 472 (Trp^{43}, TGG), resulting in pHPI-1342 and pHPI-1345, respectively for HCV-1 and HCV-1a (H). Mutation N21 changed the 79^{th} codon of the core+1 ORF at nt 580 (Gly^{79}, GGT) to a TAG stop codon, yielding pHPI-1380 (HCV-1), and pHPI-1398 (HCV-1a (H)). Finally, mutation N22 introduced a TAG terminator codon eight codons downstream of mutation N21 at nt 604 (Met^{87}), giving rise to pHPI-1381 (HCV-1) and pHPI-1397 (HCV-1a (H)). As expected, *in vitro* expression from the plasmids carrying the N1, N21, and N22 mutations failed to support the production of detectable levels of the core+1-LUC protein (Fig. 5, *A–Dc*, *lanes 2* and *3*). Again, when the same plasmids were used to analyze the production of the core+1-LUC protein in BHK-21 and Huh-7 cells, the resulting data were totally different from those obtained *in vitro*. Specifically, in contrast to the *in vitro* data, mutations N1 and N21 had no significant effect on the expression of the core+1-LUC gene as shown by

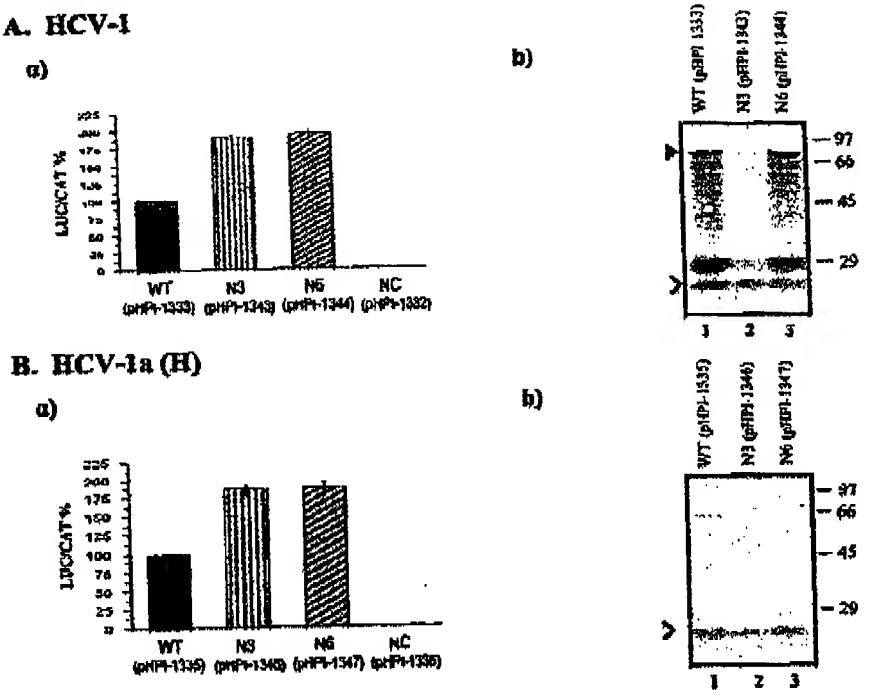

FIG. 4. **Effect of mutations within nucleotide sequences that flank codons 8–11 of the HCV-1 (*A*) and the HCV-1a (H) (*B*) core-coding region on the expression of core+1-LUC chimeric gene.** The wild-type pHPI-1333 (HCV-1) and pHPI-1335 (HCV-1a (H)) and N3, N6 mutant variants pHPI-1343, pHPI-1344 (HCV-1), and pHPI-1345, pHPI-1347 (HCV-1a (H)), respectively, were used to transfect BHK-21 cells (*a*) or transcribed *in vitro* and equal amounts of RNAs were translated in Flexi rabbit reticulocyte lysates (*b*). *a*, duplicate cultures of BHK-21 cells were transfected with the wild-type or the mutated constructs. The relative activity of each mutant variant was determined as described in the legend of Fig. 2. *Bars* represent the means from two separate experiments each performed in duplicate. *Error bars* indicate S.D. *b*, translation products were resolved by SDS-PAGE and analyzed by autoradiography. *Filled* and *open arrowheads* show the chimeric core+1-LUC and CAT proteins, respectively. *WT*, wild type; *NC*, negative control.

the fact that the level of luciferase activity was about the same as in the wild-type constructs in both BHK-21 and Huh-7 cells (Fig. 5, *A-D*, *a* and *b*). No differences were observed between HCV-1 (pHPI-1342, -1380), and HCV-1a (H) (pHPI-1345, -1398). On the contrary, the N22 mutation almost completely abolished the synthesis of the core+1-LUC protein from both HCV-1 and HCV-1a (H), in both BHK-21 and Huh-7 cell lines, as the levels of luciferase activity were similar to those of the negative control (Fig. 5, *C* and *D*, *a* and *b*). Only background levels of luciferase activity were detected from the negative controls pHPI-1332 (HCV-1) and pHPI-1336 (HCV-1a (H)). Therefore, these data support the conclusion that efficient translational initiation of the core+1 ORF in transfected cells is mediated from internal initiation codons that may be located between nt 583 and 606 (codons 80–87).

Secondly, as the region between nt 583 and 606 (codons 80–87) contains two ATGs (nt 598-ATGNNNATG-606), we assessed the functional importance of these ATGs as initiation sites for the translation of the core+1 protein *in vivo*. For this purpose, we changed ATG^{598} (85th codon) and ATG^{604} (87th codon) of the dicistronic HCV-1 construct pHPI-1333 simultaneously and separately to the non-initiator codon GGG. Mutation N25 (Fig. 3*B*) converted both methionines at positions 85 and 87 to glycines, resulting in pHPI-1401, whereas mutation N23 (Fig. 3*B*) altered only Met^{85} and mutation N24 (Fig. 3*B*) altered only Met^{87} giving rise to pHPI-1399 and pHPI-1400, respectively. The transfection of BHK-21 (*A*) and Huh-7 (*B*) with mutants pHPI-1399 (N23) and pHPI-1400 (N24) yielded similar levels of luciferase translation as the wild-type construct (Fig. 6). In contrast, mutation N25 severely affected the production of the chimeric core+1-LUC protein, which was about 23% of the wild-type level in BHK-21 cells (Fig. 6*A*) and about 26% in Huh-7 (Fig. 6*B*). These results suggest that the two methionines (Met^{85} and Met^{87}) of the core+1-coding region are involved in core+1 expression, as conversion of both of them to glycine significantly reduced the levels of luciferase activity. Interestingly, however, the conversion of each of the two methionines separately to glycine had no effect on the expression of the core+1 ORF.

Finally, to confirm the above data, we compared the size of the core+1-LUC proteins produced *in vivo* and *in vitro*. Taking into account the above results, we expected than the protein predominantly produced in transfected cells would have a molecular mass of about 62 kDa, that is 10 kDa shorter that the 72-kDa core+1-LUC protein synthesized *in vitro*. To test this hypothesis, the IRES-core+1-LUC cassette contained in the dicistronic construct pHPI-1333 (HCV-1), as well as the corresponding negative control IRES-core-1-LUC cassette of pHPI-1332 were transferred into a monocistronic expression vector under the control of a CMV promoter, resulting in pHPI-1362 and pHPI-1363 respectively (Fig. 7*A*). This system improves the detection of the luciferase protein, as HCV IRES is more active in monocistronic constructs. Specifically, the luciferase activity exhibited by the monocistronic IRES-core+1-LUC construct pHPI-1362 in BHK-21 cells forty-eight hours post-transfection was about 9-fold higher than that yielded from the dicistronic pHPI-1333 (Fig. 7*B*). Only background levels of luciferase activity were exhibited from the negative control pHPI-1363. Thus, immunoprecipitation experiments were carried out with extracts of BHK-21 cells transfected with pHPI-1362, using a goat polyclonal antibody raised against luciferase. As predicted from our mutational analysis, a protein with an apparent molecular mass of around 62 kDa reacted strongly with the polyclonal antibody (Fig. 7*C*, *lane 2*). This protein was clearly smaller than the chimeric core+1-LUC protein produced *in vitro* from the pHPI-1333 construct (Fig. 7*C*, *lane 3*). No protein was produced by the negative control pHPI-1363 (Fig. 7*C*, *lane 1*). These results are consistent with our mutagenesis data (Figs. 5 and 6) and support the hypothesis that an internal translational initiation site is being used during the core+1 synthesis *in vivo*.

Overall, these data provide strong evidence that a shorter

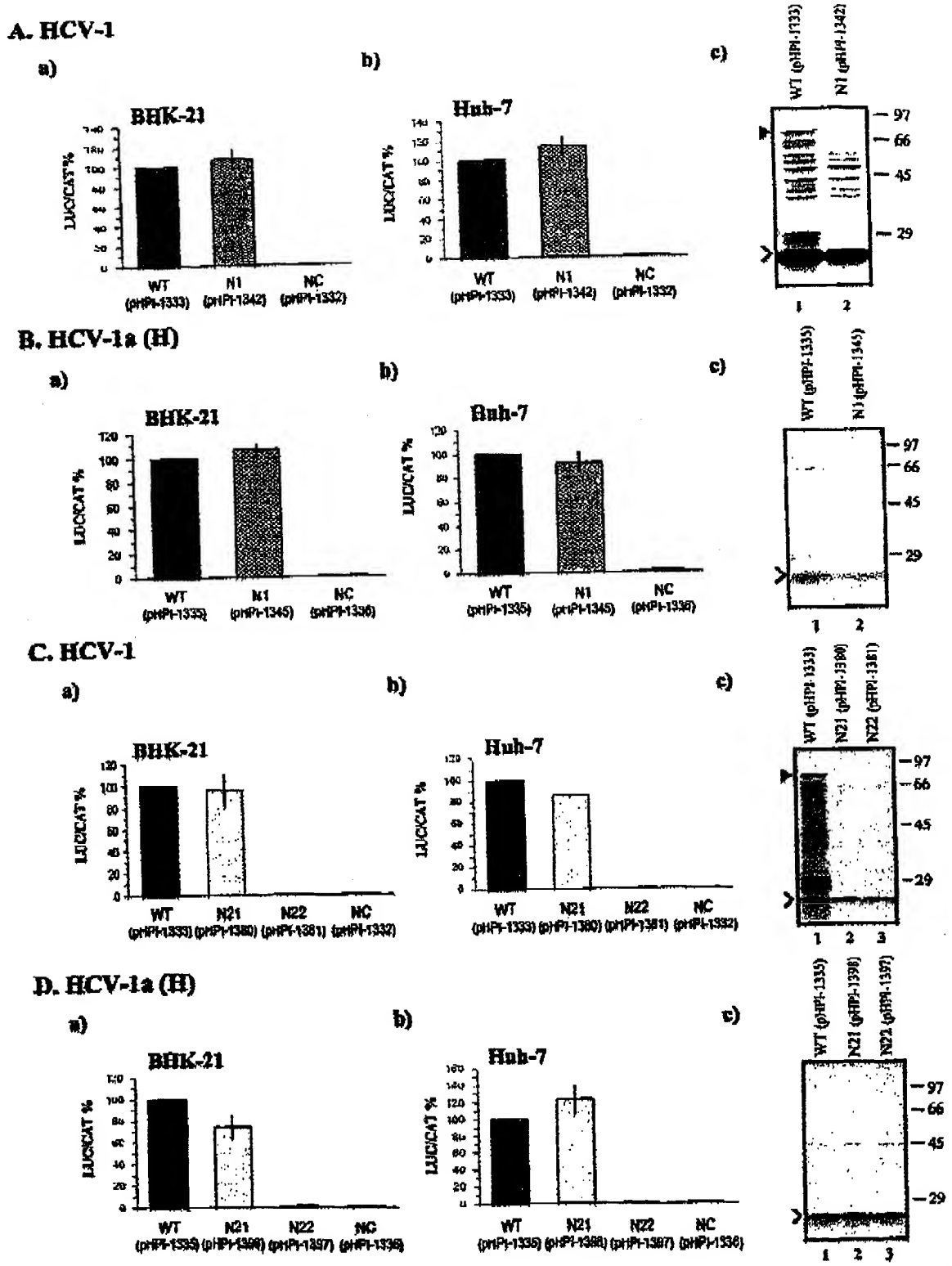

FIG. 6. **Mutational analysis within the core+1-coding sequence of HCV-1 and HCV-1a (H) isolates.** The HCV-1 (*A* and *C*) and HCV-1a (H) (*B* and *D*) wild type (pHPI-1333 and pHPI-1335, respectively), and mutated plasmids (pHPI-1342 (N1), -1380 (N21), -1381 (N22), and pHPI-1345 (N1), -1398 (N21), -397 (N22), respectively) were expressed in BHK-21 (*a*) and Huh-7 (*b*) cells or in Flexi rabbit reticulocyte lysates (*c*).

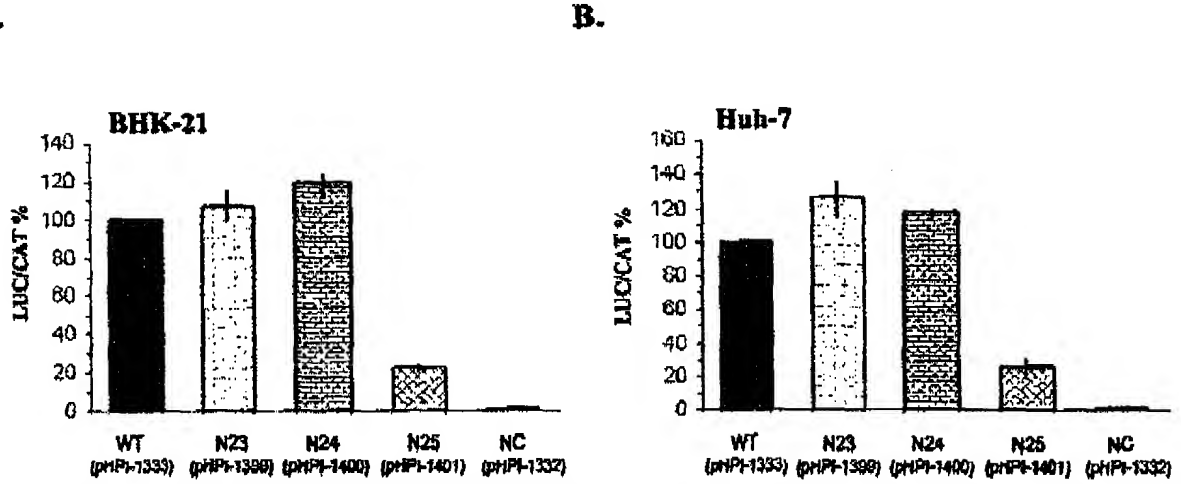

FIG. 6. **Effect of mutations targeting codons ATG^{598} and ATG^{604} of the core+1-coding sequence.** Duplicate cultures of BHK-21 (*A*) and Huh-7 (*B*) cells were transfected with the dicistronic HCV-1 wild-type (pHPI-1333) and mutated constructs: pHPI-1399 (N23), pHPI-1400 (N24), and pHPI-1401 (N25). The relative activity of each mutant variant was calculated as described in the legend of Fig. 2. *Bars* represent the means from two separate experiments each carried out in duplicate. *Error bars* indicate the S.D. *WT*, wild type, *NC*, negative control.

form of core+1 protein is produced in transfected cells by an alternative translational mechanism that directs efficient translational initiation from internal codons within the core+1-coding sequence.

DISCUSSION

In this study, we provide several lines of evidence supporting an alternative translation mechanism for the expression of the HCV core+1 ORF in transfected cells. This alternative mechanism is predicted to direct the synthesis of a shorter form of the ARFP/F/core+1 protein.

First, the core+1 ORF from both HCV-1 and HCV-1a (H) isolates is efficiently translated in transfected cells as shown by the high levels of luciferase activity produced from dicistronic constructs containing the HCV-1 or HCV-1a (H) core cDNA sequence fused with the luciferase gene in the +1 frame. This is in contrast to the results of expression studies in rabbit reticulocyte lysates (Fig. 1C*b*), which reproducibly failed to detect expression of the core+1 ORF from the HCV-1a (H) isolate (9).

Second, naturally occurring mutations identified within codons 9–11 of the core-coding sequence (N18 and N19 for HCV-1 or N15 and N16 for HCV-1a (H)) in clinical isolates from cancer patients had no effect on the translation of the core+1-LUC in transfected cells. This indicates that the expression of the core+1 ORF is independent on the A-rich nature of codons 8–11, the region proposed to be the frameshift site for the core+1 translation *in vitro*. In contrast, these mutations abolished the production of the hybrid protein in rabbit reticulocyte lysates, supporting the critical role of this region for the *in vitro* production of the ARFP/F/core+1 protein (p16/p17).

Third, the N3 mutation, which converted the initiator ATG codon of the HCV polyprotein into a stop codon, did not abolish the production of the core+1-LUC fusion protein *in vivo*, indicating that efficient translation initiation of the core+1 ORF does not require the polyprotein initiator codon. In contrast, as expected from previous studies, the same mutation failed to initiate the translation of the core+1-LUC protein *in vitro*.

Fourth, further mutational studies suggested that the translational initiation site for core+1 is located between nt 583 and 606 within the core-coding sequence, as the nonsense mutation of the 43rd codon, at nt 472 (mutation N1), or of the 79th codon, at nt 580 (mutation N21) of the core+1 ORF did not affect core+1-LUC translation, whereas the nonsense mutation of codon 87, at nt 604 (mutation N22), abolished the production of the core+1-LUC fusion protein (arbitrarily starting measurement from the GCA alanine codon of the core+1 ORF). Additionally, we showed that the efficiency of core+1 expression is dependent on the presence of ATG^{598} or/and ATG^{604}, the two ATGs contained in the region of nt 583–606, as mutation N25, which converted both of these ATGs to GGG, reduced core+1 expression to about 25% of that in the wild-type.

Fifth, immunoprecipitation analysis of the chimeric core+1-LUC protein produced in mammalian cells indicated that the immunoprecipitated protein was smaller (by about 10 kDa) than the core+1-LUC hybrid protein produced *in vitro* (Fig. 7).

These data suggest that in transfected cells, efficient translation initiation of the core+1 ORF is mediated from internal codon(s) located between nt 583 and 606, which may coincide with the ATG^{598} or/and ATG^{604} of the core+1 ORF. Consequently, the predominant form of ARFP/F/core+1 protein produced *in vivo* in these conditions should be smaller than the 16/17-kDa product synthesized *in vitro*, as it is predicted to lack the first 85 amino acids. Notably, the shorter form of the ARFP/F/core+1 protein is still a very basic protein (pI~12) and contains one of the two previously predicted hydrophobic domains in its N-terminal half.

It should be noted that the production of the ARFP/F/core+1 protein was not affected by the conversion of only one of the two ATGs at positions 598 and 604 of the core+1 ORF to GGG (mutations N23, N24), whereas it was significantly reduced by the mutation of both of these ATGs to GGG (mutation N25). This suggests that both ATGs are involved in the initiation of

a and *b*, the experiments were carried out in duplicate and repeated at least twice. The relative activity of each mutant variant was determined as described in the legend of Fig. 2. *Bars* represent the means. *Error bars* correspond to the S.D. *c*, translation products were separated by SDS-PAGE and analyzed by autoradiography. The positions of the hybrid core+1-LUC and the CAT proteins are indicated by *filled* and *open arrowheads*, respectively. *WT*, wild type; *NC*, negative control.

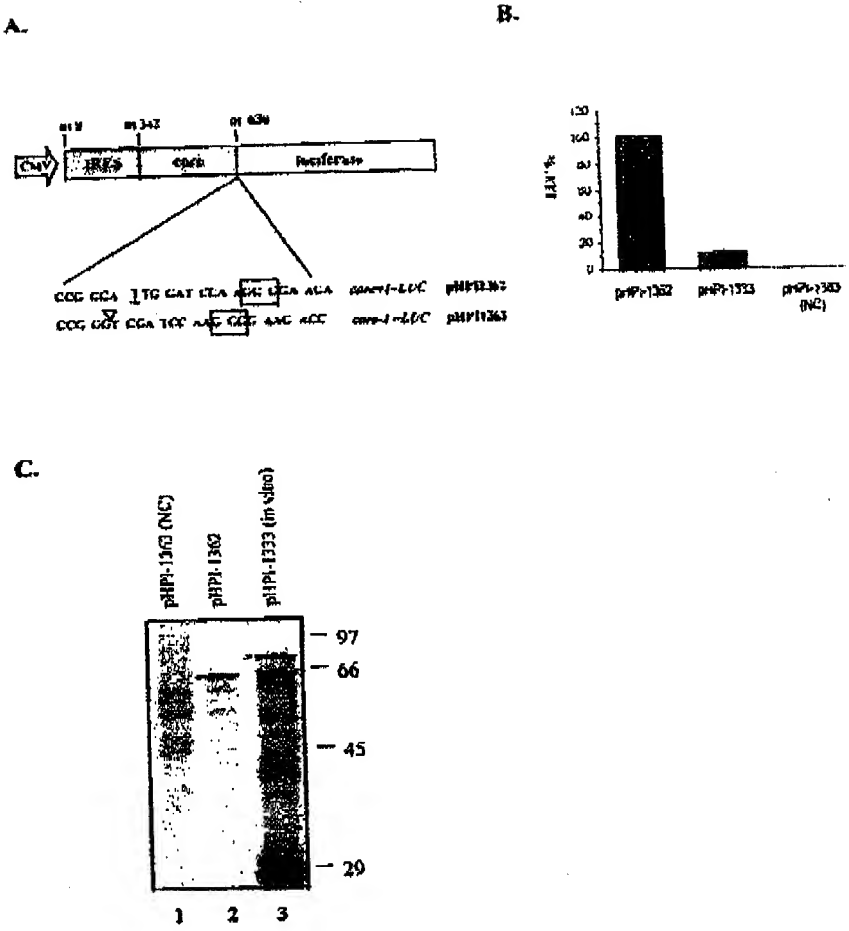

FIG. 7. **Expression of the chimeric core+1-LUC protein in transfected cells.** *Panel A*, schematic representation of the monocistronic constructs pHPI-1362 (core+1-LUC) and pHPI-1363 (core-1-LUC). *Panel B*, duplicate cultures of BHK-21 cells were transfected with the monocistronic core+1-LUC construct pHPI-1362 or the dicistronic core+1-LUC pHPI-1333 and the relative luciferase activity was determined. *Bars* represent the means from two separate experiments. *Error bars* represent the S.D. *Panel C*, immunoprecipitation of [^{35}S]methionine-labeled translation products of the core+1-LUC and core-1-LUC containing monocistronic constructs from transiently transfected BHK-21 cells using an anti-LUC goat polyclonal antibody. The immunoprecipitates were analyzed by SDS-PAGE, followed by autoradiography. The hybrid core+1-LUC protein produced *in vivo* is marked by a dot. The *open arrowhead* shows the [^{35}S]methionine-labeled core+1-LUC protein synthesized in rabbit reticulocyte lysates. *NC*, negative control.

core+1 translation and are able to substitute for each other. It is noteworthy that the ATG^{598} and ATG^{604} codons are well conserved among HCV isolates. A comparative analysis of 117 HCV core sequences from the genotypes 1a, 1b, 2a-c, 2k, 3a, 3b, 4a-f, 5a, 6a, 6d, 6e, 6h, and 6k, from the GenBank™ data base, revealed that 66 variants contain both ATG^{598} and ATG^{604}, 23 carry only the ATG^{598} or the ATG^{604}, and 5 lack both ATGs. This high level of conservation is consistent with these ATGs having a functional role during core+1 translational initiation. Notably, the dicistronic constructs pHPI-1312 (HCV-1) and pHPI-1330 (HCV-1a (H)), which contain the LUC gene fused in the −1 frame of the core-coding sequence at nt 630, displayed high levels of luciferase activity when the initiator ATG codon of the LUC gene was included (data not shown). This expression was abolished only after the conversion of the ATG start codon of LUC to GGG. This finding suggests that any ATG start codon located in the vicinity of nt 583 and 606 is functional.

It should also be emphasized that the conversion of both ATGs to GGG did not completely abolish the expression of the core+1-LUC protein (25% of the wild-type level), which is consistent with the concomitant expression of the larger (16/17 kDa) form of the core+1 protein. Indeed, we have preliminary evidence supporting the expression of both forms of the ARFP/F/core+1 protein (data not shown). Clearly, however, in the conditions used in our experiments translation of the core+1 ORF from internal initiation codons was very efficient.

The physiological implications of these findings on the expression of core+1 *in vivo* remain to be elucidated. However, the short form of the ARFP/F/core+1 protein is probably synthesized in conditions that restrict the expression of the viral polyprotein, inasmuch as suppression of core expression in the presence of mutation N3 or N6 failed to abolish the production of core+1-LUC. This suggests that the ARFP/F/core+1 protein may play a critical role in controlling the life cycle of the virus.

RNA and to a lesser extent DNA viruses that are subjected to genome size constraints have devised strategies to expand their coding capacity, such as ribosomal frameshifting (16) and internal translational initiation. Different mechanisms that allow escape from an upstream initiator codon and direct initiation from internal codons have been described; these include context-dependent leaky-scanning (13, 17, 18, 19, 20, 21), ribosome shunting (17, 22, 23, 24, 25, 26, 27), and IRES-mediated initiation (28). The internal initiation of ARFP/F/core+1 translation *in vivo* might include features of one of these mechanisms. If the expression of core+1 is indeed mediated by an IRES element, this element is predicted to be a different IRES from that located at the 5′-end of the viral RNA, as we have data suggesting that core+1 translation *in vivo* does not require the HCV IRES (data not shown).

The reason for the divergence between our *in vitro* and *in*

vivo results for the principal mechanism involved in the core+1 translation remains unclear. Recent data have shown that another mechanism related to *in vivo* translational initiation, the interaction between the cap and poly(A) tail of eukaryotic mRNAs, does not occur in Flexi rabbit reticulocyte lysates. Instead, the mechanism operates in a modified system based on ribosome-depleted rabbit reticulocyte lysates (29, 30). Alternatively, the translation of the ARFP/F/core+1 protein may be regulated by different mechanisms depending on the cellular conditions and the ways in which the translation machinery is modified during HCV infection.

Acknowledgments—Plasmids path 10/17–38 and path 11/38–27 were obtained from the Centers for Disease Control and Prevention, National Center for Infectious Diseases Plasmid pDNA-C1 was kindly provided by Dr. G Inchauspe We thank N. Michaelidou for excellent technical assistance with tissue culture and our colleagues from the Molecular Virology laboratory for helpful discussions.

REFERENCES

1. Chuo, Q. L., Kuo, G., Weiner, A. J., Overby, L. R., Bradley, D. W., and Houghton, M. (1989) *Science* 244, 359–362
2. Saito, I., Miyamura, T., Ohbayashi, A., Harada, H., Katayama, T., Kikuchi, S., Watanabe, Y., Koi, S., Onji, M., Ohta, Y., *et al* (1990) *Proc. Natl. Acad. Sci. U. S. A.* 87, 6547–6549
3. Di Bisceglie, A. M. (2000) *Hepatology* 31, 1014–1018
4. Murphy, F. A., Fauquet, C. M., Bishop, D. H. L., Ghabrial, S. A., Jarvis, A. W., Martelli, G. P., Mayo, M. A., and Summers, M. D. (eds) (1995) *Virus Taxonomy. Sixth Report of the International Committee on Taxonomy of Viruses*, pp. 424–426, Springer-Verlag, Vienna & New York
5. Reed, K. E., and Rice, C. M. (2000) *Curr. Top. Microbiol. Immunol.* 242, 55–84
6. Rijnbrand, R. C., and Lemon, S. M. (2000) *Curr. Top. Microbiol. Immunol.* 242, 85–116
7. Ray, R. B., and Ray, R. (2001) *FEMS Microbiol. Lett.* 202, 149–156
8. McLauchlan, J. (2000) *J. Viral Hepat.* 7, 2–14
9. Varaklioti, A., Vassilaki, N., Georgopoulou, U., and Mavromara, P. (2002) *J. Biol. Chem.* 277, 17713–17721
10. Walewski, J. L., Keller, T. R., Stump, D. D., and Branch, A. D. (2001) *RNA* 7, 710–721
11. Xu, Z., Choi, J., Yen, T. S., Lu, W., Strohecker, A., Govindarajan, S., Chien, D., Selby, M. J., and Ou, J. (2001) *EMBO J.* 20, 3840–3848
12. Rumenapf, T., Stark, R., Heimann, M., and Thiel, H. J. (1998) *J. Virol.* 72, 2544–2547
13. Yamasaki, K., Wohl, C. C., and Ross, R. P. (1999) *J. Virol.* 73, 8519–8526
14. Pavlaki, I., Georgopoulou, U., Varaklioti, A., and Mavromara, P. (1999) *FEBS Lett.* 453, 49–53
15. Yeh, C. T., Lo, S. Y., Dai, D. I., Tang, J. H., Chu, C. M., and Liaw, Y. F. (2000) *J. Gastroenterol. Hepatol.* 15, 182–191
16. Brierley, I. (1995) *J. Gen. Virol.* 76, 1885–1892
17. Lavorre, P., Kolakofsky, D., and Curran, J. (1998) *Mol. Cell. Biol.* 18, 5021–5031
18. Kobayashi, T., Watanabe, M., Kamohara, W., Tomonaga, K., and Ikuta, K. (2000) *Virology* 277, 296–305
19. Stacey, S. N., Jordan, D., Williamson, A. J., Brown, M., Coote, J. H., and Arrand, J. R. (2000) *J. Virol.* 74, 7284–7297
20. Pavlakis, G. N., and Felber, B. K. (1990) *New Biol.* 2, 20–31
21. Jayaraj, H. B., and Whitt, M. A. (2002) *J. Virol.* 76, 8011–8018
22. Curran, J., and Kolakofsky, D. (1988) *EMBO J.* 7, 2869–2874
23. Futterer, J., Kiss-Laszlo, Z., and Hohn, T. (1993) *Cell* 73, 789–802
24. Ryabova, L. A., Pooggin, M. M., Domingues, D. L., and Hohn, T. (2000) *J. Biol. Chem.* 275, 37278–37284
25. Futterer, J., Potrykus, I., Bao, Y., Li, L., Burns, T. M., Hull, R., and Hohn, T. (1996) *J. Virol.* 70, 2999–3010
26. Yueh, A., and Schneider, R. J. (2000) *Genes Dev.* 14, 414–421
27. Reuter, M., Reuter, A., and Unsay, M. (1999) *J. Virol.* 73, 3062–3070
28. Grundhoff, A., and Ganem, D. (2001) *J. Virol.* 75, 1857–1863
29. Michel, Y. M., Poncet, D., Piron, M., Kean, K. M., and Borman, A. M. (2000) *J. Biol. Chem.* 275, 32268–32276
30. Borman, A. M., Michel, Y. M., and Kean, K. M. (2000) *Nucleic Acids Res.* 28, 4068–4075

PTO/SB/17 (10-03)
Approved for use through 7/31/2006. OMB 0651-0032
U.S. Patent and Trademark Office: U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

FEE TRANSMITTAL for FY 2004

Effective 10/01/2003. Patent fees are subject to annual revision.

[X] Applicant claims small entity status. See 37 CFR 1.27

TOTAL AMOUNT OF PAYMENT ($) 55.00

Complete if Known	
Application Number	09/719,277-Conf. #7039
Filing Date	April 13, 2001
First Named Inventor	Andrea D. Branch
Examiner Name	D. Wortman
Art Unit	1648
Attorney Docket No	RII-003CPUS

METHOD OF PAYMENT (check all that apply)

[ ] Check [ ] Credit Card [ ] Money Order [ ] Other [ ] None

[X] Deposit Account:

Deposit Account Number: 12-0080

Deposit Account Name: Lahive & Cockfield, LLP

The Director is authorized to: (check all that apply)

[X] Charge fee(s) indicated below [X] Credit any overpayments

[X] Charge any additional fee(s) or any underpayment of fee(s)

[ ] Charge fee(s) indicated below, except for the filing fee to the above-identified deposit account.

FEE CALCULATION

1. BASIC FILING FEE

Large Entity Fee Code	Large Entity Fee ($)	Small Entity Fee Code	Small Entity Fee ($)	Fee Description	Fee Paid
1001	770	2001	385	Utility filing fee	
1002	340	2002	170	Design filing fee	
1003	530	2003	265	Plant filing fee	
1004	770	2004	385	Reissue filing fee	
1005	160	2005	80	Provisional filing fee	

SUBTOTAL (1) ($) 0.00

2. EXTRA CLAIM FEES FOR UTILITY AND REISSUE

			Extra Claims		Fee from below		Fee Paid
Total Claims	36	-55** =		x		=	0.00
Independent Claims	1	-7** =		x		=	0.00
Multiple Dependent						=	

Large Entity Fee Code	Large Entity Fee ($)	Small Entity Fee Code	Small Entity Fee ($)	Fee Description
1202	18	2202	9	Claims in excess of 20
1201	86	2201	43	Independent claims in excess of 3
1203	290	2203	145	Multiple dependent claim, if not paid
1204	86	2204	43	** Reissue independent claims over original patent
1205	18	2205	9	** Reissue claims in excess of 20 and over original patent

SUBTOTAL (2) ($) 0.00

**or number previously paid, if greater; For Reissues, see above

FEE CALCULATION (continued)

3. ADDITIONAL FEES

Large Entity Fee Code	Large Entity Fee ($)	Small Entity Fee Code	Small Entity Fee ($)	Fee Description	Fee Paid
1051	130	2051	65	Surcharge – late filing fee or oath	
1052	50	2052	25	Surcharge – late provisional filing fee or cover sheet.	
1053	130	1053	130	Non-English specification	
1812	2,520	1812	2,520	For filing a request for ex parte reexamination	
1804	920*	1804	920*	Requesting publication of SIR prior to Examiner action	
1805	1,840*	1805	1,840*	Requesting publication of SIR after Examiner action	
1251	110	2251	55	Extension for reply within first month	55.00
1252	420	2252	210	Extension for reply within second month	
1253	950	2253	475	Extension for reply within third month	
1254	1,480	2254	740	Extension for reply within fourth month	
1255	2,010	2255	1,005	Extension for reply within fifth month	
1401	330	2401	165	Notice of Appeal	
1402	330	2402	165	Filing a brief in support of an appeal	
1403	290	2403	145	Request for oral hearing	
1451	1,510	1451	1,510	Petition to institute a public use proceeding	
1452	110	2452	55	Petition to revive – unavoidable	
1453	1,330	2453	665	Petition to revive – unintentional	
1501	1,330	2501	665	Utility issue fee (or reissue)	
1502	480	2502	240	Design issue fee	
1503	640	2503	320	Plant issue fee	
1460	130	1460	130	Petitions to the Commissioner	
1807	50	1807	50	Processing fee under 37 CFR 1.17(q)	
1806	180	1806	180	Submission of Information Disclosure Stmt	
8021	40	8021	40	Recording each patent assignment per property (times number of properties)	
1809	770	2809	385	Filing a submission after final rejection (37 CFR 1.129(a))	
1810	770	2810	385	For each additional invention to be examined (37CFR 1.129(b))	
1801	770	2801	385	Request for Continued Examination (RCE)	
1802	900	1802	900	Request for expedited examination of a design application	
Other fee (specify)					

*Reduced by Basic Filing Fee Paid SUBTOTAL (3) ($) 55.00

SUBMITTED BY				(Complete if applicable)	
Name (Print/Type)	Megan E. Williams	Registration No (Attorney/Agent)	43,270	Telephone	(617) 227-7400
Signature	Megan E. Williams			Date	January 5, 2004

I hereby certify that this correspondence is being facsimile transmitted to the Patent and Trademark Office, facsimile no. (703) 872-9306, on the date shown below.

Dated: January 5, 2004 Signature: Megan E. Williams (Megan E. Williams)

PTO/SB/97 (12-97)
Approved for use through 9/30/00. OMB 0651-0031
Patent and Trademark Office, U.S. DEPARTMENT OF COMMERCE

Certificate of Transmission Under 37 CFR 1.8

I hereby certify that this correspondence is being facsimile transmitted to the United States Patent and Trademark Office

on January 5, 2004
Date

Signature

Megan E. Williams
Typed or printed name of person signing Certificate

Note: Each paper must have its own certificate of transmission, or this certificate must identify each submitted paper.

One Month Request for Extension of Time Under 37 CFR 1.136(a) (1 page);
Amendment Transmittal (1 page);
Amendment and Response with attached Appendices A-D (56 pages);
Fee Transmittal (1 page in duplicate);
Certificate of Transmission under 37 CFR 1.8 (1 page); and
Charge $55.00 to deposit account 12-0080.